IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application No.	:	10/629,347	Confirmation No.:	6349
Applicant	:	Ciccarelli		
Title	:	Systems and Methods for Providing True Scale Measurements for Digitized Drawings		
Filed	:	July 28, 2003		
Art Unit	:	2625		
Examiner	:	Thierry L. Pham		
Attorney Docket No.	:	ARC11.012		

DECLARATION OF LAWRENCE A. WHITE PURSUANT TO 37 C.F.R. § 1.132

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

I, Lawrence A. White, hereby declare that:

I.

II. INTRODUCTION

1. I am a resident of Duluth, GA and have been working in the field of Information Technology for 25 years.

2. A copy of my Curriculum Vitae is attached hereto as Exhibit A. As evident from my Curriculum Vitae, I am fully familiar with computer-aided design (CAD) programs and products. Also, in my career I have designed and implemented Emergency and Disaster Management Centers, both domestically and internationally. A sampling of these centers and emergency management agencies include Department of Homeland Security, Urban Area Security Initiatives (DHS UASI), St. Louis Area Regional Response System (STARRS), New Jersey State Police, State of Missouri Emergency Management Agency and Gauteng (South Africa) Provincial Disaster Management Centre.

3. I am currently the Chief Technology Officer ("CTO") of Archaio, LLC ("Archaio"), the assignee of the present patent application. I have held this position since January 2009. As the CTO of Archaio, I am familiar with the various products that have been developed and that are currently under development. I am also familiar with the marketing and sales of the various products that have been developed.

4. I am familiar with U.S. Pat. App. No. 10/629,347, entitled "Systems and Methods for Providing True Scale Measurements for Digitized Drawings," hereinafter referred to as the Patent Application. I am also familiar with the development of the various products set forth in the Patent Application. The design and development process, as well as the later commercial success, of the various products described in the Patent Application are briefly described below. In particular, the considerations for embedding scale information in the header of a raster file are discussed.

III. MATERIALS CONSIDERED

5. I have reviewed and am familiar with the office action dated February 17, 2009 in the Patent Application and the prior art references U.S. Patent No. 6,134,338 to Solberg ("Solberg") and with U.S. Patent Publication No. 2002/0077787 to Rappaport, et al. ("Rappaport") applied by the Office against the claims of the Patent Application. I have also reviewed and am familiar with the set of amended claims submitted to the U.S.P.T.O. on August 17, 2009.

IV. PROCESS FOR DETERMINING WHETHER CLAIMS ARE OBVIOUS

6. Having read the office action and the prior art references I have formulated the opinions provided below regarding obviousness.

V. ORDINARY SKILL IN THE ART

7. I believe the hypothetical person of ordinary skill in the art during the 2003 timeframe would have been a person having an undergraduate degree in computer science and post graduate studies in management information systems that has practical or academic experience in the field of computer programming. In the 2003 timeframe I was employed with IBM Corporation as a Global Services Business Consulting Principle responsible for sizing,

scoping, and delivering advanced technology services to the U.S. Federal Marketplace, with a focus on enterprise system management and emergency response systems for command and control for military and intelligence agencies.

VI. Overview of the Patent Application

8. It is estimated that approximately 85% of the world's infrastructure plans, such as architectural, mechanical, structural, and electrical diagrams, are currently stored in paper format only. The digitizing of paper infrastructure plans into a raster image, known as raster files, are recognized for providing the benefit of being good for archiving, printing and sharing. (See Exhibit B and Exhibit C.) However, despite the benefits of raster files and the fact that raster files are non-proprietary, the industry largely only utilizes raster files for archiving or transferring because making considerable modifications to raster files are found to be tedious. (See Exhibit C.)

9. At and around the time of the creation of the present Patent Application, the industry concentrated on converting or transforming raster files for use with computer-aided design (CAD) programs as CAD has the known ability to precisely describe, create, scale and manipulate individual objects. (See Exhibit C and Exhibit D "[v]ector files can be scaled, which means one can zoom in on the details of a drawing. Also, they are more easily edited than raster files.")

10. Again despite the benefits of raster files, the industry clearly diverged or taught away from using raster files for reasons other than archiving, printing, sharing or making minor modifications. (See for instance Exhibit C and Exhibit B which finds that "[r]aster images can only be edited by adjusting the values of individual dots.") The industry deemed that vectors could be made to be mathematically perfect, while rasters could not. (See Exhibit C.) Thus, CAD was understood to be the standard in technical drafting notwithstanding the fact that CAD programs typically utilize their own propriety format and therefore are often non-transferable between different systems and programs. (See Exhibit H.) Another problem with CAD programs, such as and including Solberg, is the fact that scale information is stored in external library files. If the library file is not available or if the library file becomes corrupted, the scale information may be unavailable when a raster file is accessed.

11. The inventor of the Patent Application however recognized a clear, unaddressed, and long felt need for developing non-proprietary systems and methods to electronically store infrastructure plans in an accurate, scaled and secure manner, and share infrastructure plans with data integrity.

12. Traditionally when paper plans are scanned and digitized for electronic storage, the images original physical size, and therefore the corresponding usefulness of the image scale, of a particular document is no longer a concrete attribute of the image. For example, if a paper version of an infrastructure plan is thirty inches in height and forty inches in width and then scanned, a computer user of that scanned electronic image would see the document as a different physical size when using different monitors depending on the size of the display device and its own pixel resolution. Thus, the scale that appears on the document (e.g., one eighth inch equals one foot, etc.) will be incorrect when an electronic depiction of the document is displayed on a computer monitor. This is because the original physical size of a paper image has no direct correlation to the pixel dimensions of a computer monitor. As a result, a 20 inch wide monitor can only display an image as twenty inches wide if viewing the whole image and a twenty-five inch wide monitor can only display an image as twenty-five inches wide if viewing the whole image. Also, neither monitor would be able to display the whole image as it originally appeared, that is, as a forty inch wide image. The user has no way to know what the original physical size of the paper drawing was, yet the scale ratio of the image listed on the plan is directly tied to the physical size of the original paper document. So if a computer user viewing the scanned infrastructure plan on a twenty-five inch monitor tried to take a physical measurement of the image on the computer monitor using that data with the image scale to manually compute a true scale measurement the result would be a wrong measurement value. Furthermore zooming the image so that only portions of the original image appear on the computer monitor also distorts the physical size of the image making any physical measurement of an image or image element not useful when combined with scale to calculate a true scale dimension measurement. In essence, once a paper drawing is scanned, the scale information on that drawing is no longer valid and accurate when a digital version of the paper drawing is viewed on a monitor or display device.

13. In response to the long felt need for systems and methods to electronically store infrastructure plans while maintaining true and accurate scale information and the need to permanently secure scale to digitized plans, the inventor of the Patent Application embeds scale information in a header of a digitized raster image of an infrastructure plan. The inventor of the Patent Application employs a specific private Tagged Information File Format (TIFF) header tag from Adobe Systems to store scale information.

14. This private header tag is a dedicated location that permits scale information to be secured indefinitely within a header of a raster image. As a result of embedding the scale information within the header of a raster file, the digital raster image and all the data needed to calculate dimension measurement data at a future time can be stored as a single file. Also, as there are hundreds and thousands of tags the embedding of scale information in the dedicated location permits quick and easy access to such scale information rather than having to unnecessarily delve through the thousands of tags in the TIFF header.

15. The act of embedding scale information within a header of the digital raster image, using the dedicated private header tags and storing the digital raster image as a single file are important aspects of the Patent Application.

16. The Patent Application, by way of these important aspects, provides a raster file that secures the scale information to the single digital raster image as it is known that integrity of a single raster file can better maintained than that of multiple raster files, which easily disassociate from one another. Also, the Patent Application prevents scale information that is embedded in the header of a raster image from being lost or overwritten as it is known that the information stored in a raster header location is lost or overwritten much less than the data kept in the main body of a file. Further, raster images are by their nature not able to store data other than pixel data in the main body of the electronic file.

17. The file of the Patent Application can be easily transferred between different systems and software programs and stored to be readily available for subsequent access.

18. The file of the Patent Application can also be quickly opened and the embedded scale information quickly accessed in order to determine true scale measurement information upon subsequent access of the digitized raster file.

19. For example, an architectural drawing may be converted to a digitized raster file and scale information can be embedded in the header of the digitized raster file. Once the digitized raster file is rendered, a user may draw a line or shape in the rendered architectural drawing, and true scale measurements for the drawn line or shape (e.g., distance or area) may be determined utilizing the embedded scale information.

20. By creating the single digital raster image file with scale embedded in the dedicated location of the header of the single file, the inventor of the Patent Application provides a non-proprietary electronic file format that is readily available for use by a wide variety of different individuals. Users of the Patent Application are not required to be familiar with sophisticated software programs and products, such as computer-aided design (CAD) programs and products. Thus, unskilled users may be able to quickly and efficiently utilize the claimed invention with ease.

21. A wide variety of dimension data can be calculated using only the raster image with the scale data embedded in the raster image header. By using the present Patent Application, measurements can be calculated from input that is not previously prescribed. For instance, a line can be drawn on a raster image from the middle of a wall in a room to any other point in the room and a true scale measurement of the drawn line can be obtained using the scale embedded in the dedicated location of the digital raster image. These non-prescribed measurement values are not listed on the original image as it would be impossible to list all possible element measurements and combination of element measurements on a paper drawing image.

22. Also, a wide variety of different applications may utilize the systems and methods set forth in the Patent Application. Some example applications include Emergency Management and Response applications. (See Co-pending Application No. 11/068,268). Raster images of building schematics may be utilized to calculate true scale measurements for the movement of emergency personnel within the building. Given the location of emergency personnel (e.g.,

firefighters, etc.) within the building, a route for the personnel and true scale measurements associated with the route may be determined. The emergency personnel may then be provided with accurate directions while inside the building.

VII. THE SOLBERG REFERENCE ALONE OR COMBINED WITH THE RAPPAPORT REFERENCE DOES NOT RENDER OBVIOUS CLAIMS 1-2, 4-20.

A. The Solberg Patent Fails To Teach Claimed Elements

23. Solberg discloses a CAD program that utilizes raster files as intermediate files in the creation of a complex CAD file. (See Solberg Col. 22, lines 63-65). Specifically, Solberg scans a paper document and stores pieces of the original paper document as multiple raster files that are then transformed into a CAD file, a second, associated, but clearly distinguishable electronic file entity. (See Fig. 2, Steps 3.3 and 3.4, Col. 25, lines 3-7 and Fig. 4 Col. 26, lines 15-25.) Solberg's invention is designed to automate the transfer of 2D raster image data concerning real world objects into mathematically accurate 3D vector models. (See Solberg Abstract and Col. 40 lines 6-14).

24. I agree with the statement in the office action finding that Solberg does not teach embedding scale in the header of a digital raster image. As mentioned above, CAD programs, such as and including Solberg, preserve scale information in an external library file, which if corrupted or separated from the raster file will result in the loss of the scale information. Scale information may also be read from a scanned document utilizing optical character recognition techniques or keyed in upon prompting, however again the scale information is <u>not</u> embedded in a raster image. See Solberg col. 14, lines 25-54, col. 19 line 45 to col. 20, line 8 and Col. 16 lines 36- col. 17 line 16). Furthermore, in view of the set of amended claims submitted on August 17, 2009, I believe that Solberg does not teach embedding scale in a dedicated location of a header of a digital raster image.

25. Solberg also does not teach storing the digital raster image as a single file with scale embedded in the header of the single digital raster image file as presented in the set of amended claims submitted on August 17, 2009. The Patent Application preserves the original raster bitmap format as a single document and simultaneously preserves scale. In contrast,

Solberg teaches the creation of a plurality of viewpoints raster files and discloses creating a separate raster file, a floating viewpoint 242, which contains scale 187. (Solberg col. 18 line 61-col. 20 line 14.) This scale information is a separate and distinct raster file from the raster file(s) used to store the plurality of viewpoints 122. (See col. 14, lines 25-54 and col. 19 line 45 to col. 20, line 8).

26. Solberg fails to teach associating scale with raster files, rather Solberg teaches associating scale with a CAD file. After a CAD viewport is selected in steps 3.1-3.2, the AUTOCAD program prompts the user for the drawing scale in step 3.3 and then later the raster file is imported into the AUTOCAD program (See Solberg Col. 25, lines 9-14). Clearly the scale has not been set for the raster file as it had yet to be imported. Scale may also be read from the alphanumeric text representing scale information on the face of the raster image once in the CAD program. (See Solberg at Col. 16 lines 36- col. 17 line 16.) Thus, Solberg is not associating scale with a raster image and not embedding scale in a dedicated location of a header of a single raster image.

27. Furthermore, Solberg does not teach calculating a true scale measurement of a drawn line or shape based at least in part on the embedded scale information in said dedicated location of said header of said single file. Solberg instead teaches converting hard copy drawings into mathematically accurate vectors corresponding to physical dimension and edges of 3D objects and moiety that symbols represent. (Abstract, col. 13 lines 40-45). The result of Solberg is to use the vector file as a 3D computer model of the 3D object and the moiety represented by a symbol. (Col. 40, lines 6-14.) Also, the mathematically accurate AUTOCAD drawing file may be printed to create a new hard copy of the newly created converted engineering drawing. (Solberg col. 57 lines 1-15.)

28. Solberg's creation of mathematically accurate vectors traced over a previously drawn line does not teach the calculation of a true scale measurement of drawing input. (See Solberg Step 6.) Solberg simply correlates dimension information to a shape. This dimension data that is associated with a line is accessed from the CAD library file for display. Thus, Solberg is neither accessing scale data from a header of a raster image nor utilizing scale data stored in the header to calculate a true scale measurement of a drawn line or shape. It is

understood that the measurements displayed by Solberg are limited to the prescribed dimensions shown on the face of the image.

B. The Rappaport Reference Fails To Teach Claimed Elements

29. Rappaport looks to provide an admittedly non-scaled contextual map for the association of external device or tool collected metric data wherein the metric data may be visually interpreted and associated against the image back drop of the spatial environment it describes. (Rappaport abstract, [0025] and [0075].) Rappaport is designed as a means to store measured network performance where a measurement reading is associated with some textual or graphical identifier to enable easy inspection or analysis of data by anyone especially by a less-technical or untrained individual. (Rappaport [0069].)

30. The office action asserts that Rappaport teaches embedding <u>information</u> in a header of a digital raster image. Although Rappaport teaches the known concept of storing file-identifying information in the header of a file such information is metadata or generic information. Metadata is descriptive data about data retained within the body of the computer file. (See Exhibits E, F and G.) It is a key component of data lineage as it provides basic information about the source and derivation of a data set. (See Exhibit F.) Metadata includes information that describes file content such as a file name or a file type; quality; condition and other appropriate characteristics of the data. Metadata provides the necessary information for an application to "recognize" and "understand" the file, see Exhibit E, and also, this type of information is used "to properly transmit" a file. (Rappaport [0095] and [0010].)

31. In contrast to the known concept of inserting file-identifying information (file name, date, comments, etc.) in the header, the Patent Application embeds specific image information, scale information, in a dedicated location of a header of a digital raster image that unlocks additional functional possibilities.

32. Also, Rappaport does not teach embedding scale in a digital raster file. Rappaport teaches putting information in the header of a generic computer file. The difference between a header of a generic computer file and a header of a digital raster image are significant and well-known. A header of a generic computer file has a fixed role with limited capacity. A generic file

header has 5-10 slots/fields and is used to hold limited information such as metadata. The file header of a generic computer file is not accessed for reasons other than obtaining transmission information. (See Rappaport [0010], [0095], [0116] and [0124].) Also, a generic electronic file header can only be read and properly processed if it is in a proprietary or standard file exchange format. Rappaport does not teach a generic electronic file header that is in a proprietary or standard file exchange format.

33. In contrast, the Patent Application employs the header of a raster file. The digital raster header is known to have an exponentially larger capacity than the limited capacity of a generic file header to hold information. A header of the digital raster image can have an unlimited number of Private tiff Tags that can be expanded all the way up to the TIFF file size limit of 4 GB at 8 bytes per tag. Also, computer systems are designed to intelligently process the information retained in the header of the TIFF file, a non-proprietary interchange format. This format can be universally read and processed without limitation. The act of embedding scale information in a file header securely stores the scale information in the raster file. Thus, Rappaport's teaching of placing metadata in a header of a generic computer file does not equate to or render obvious the embedding of scale in a header of a digital raster image.

34. As for the assertion in the office action that scale could be put in the header via the notes subsection, this is not viable. It is known that the notes subsection of the header has limited capacity and is used for general purposes such as to retain metadata and other limited information. (See Rappaport [0116] and [0124].) Rappaport clearly teaches using the "notes" to retain metadata such as comments. See for example the "notes" line in Fig. 3 of Rappaport which states "Notes: Location file for Blacksburg Office." This comment placed in the notes subsection of Fig. 3 is nothing more than metadata. Furthermore, a computer system cannot intelligently process information from the notes subsection of a computer file; rather the computer only is able to consume information from the notes subsection as general comments or notes.

35. Rappaport's teaching of inserting metadata or file-identifying information in the notes subsection of a header does not render obvious the embedding of scale information in a dedicated location. A dedicated location in the header of a digital raster image has an address and a location that retains specified information. A computer can intelligently process information

stored in a dedicated location. The insertion of comments in the notes subsection is not equal to embedding scale in a dedicated location. Thus, it would not be obvious to insert scale information in the notes subsection of a header as information in the notes subsection is ignored and not processed by a computer.

36. For the reasons stated above the combination of Solberg with Rappaport would not be feasible and thus would not render obvious the invention of the Patent Application as claimed.

C. The Claimed Combination Of Rappaport With Solberg Changes The Principle Operation Of Solberg, The Primary Reference, And Renders Solberg Inoperable For Its Intended Purpose

37. The office action finds it would have been obvious to combine Solberg's teaching of manually entering scale with Rappaport's teaching of placing information in the notes subsection. Specifically, the office action asserts scale could be placed in the notes subsection of a header of a digital raster image. However, such a combination would change Solberg's principle operation and render Solberg inoperable for producing mathematically accurate 3D vectors. (Solberg-Abstract).

38. Solberg sets scale for the CAD system and this act is accomplished in Step 3.3 ("Set drawing scale"). (Solberg Col. 25 lines 1-14.) The office action's suggested combination would in effect be replacing Solberg's scale setting step (Step 3.3), as scale would now be set using Rappaport. This combination would result in Solberg's CAD system not being able to read the scale as it is known that CAD systems consume information from the body of the file and not from the header of a file. Thus, the placement of scale in the header of a file would change the principle operation of Solberg as its Step 3.3 would be replaced. As a result, Solberg would be rendered inoperable for its intended purpose of attaining mathematically accurate 3D vectors.

39. The replacement of Solberg's scale setting step with Rappaport's teaching would further change Solberg's principle operation as Solberg's CAD system would not be able to read scale from the header of a file and thus would have to resort to a default scale. It is known that

CAD programs use a default scale of a 1:1 ratio, stored in their own proprietary data store format, to perform mathematical calculations and measurements. The presence of a default scale in CAD programs was demonstrated in the Personal Examiner interview held on July 20, 2009. In the CAD demonstration, a true scale measurement of a drawn line was not attained as the CAD program used a default scale. Again, Solberg would not achieve its intended purpose of attaining mathematically accurate vectors.

40. The office action's suggested combination of placing scale in the header of the digital raster image would further change Solberg's principle operation as Solberg would no longer be setting scale for the CAD application. Solberg's Step 3.3 of setting scale in the CAD system is an important and critical step as it is required in order to properly introduce the viewport raster file 124 into the CAD viewport 290. Step 3.3 is also important because the viewpoint raster file image 350 in the viewport 290 serves as a backdrop to the production of the vectors 268. (Solberg Col. 25 line 25- Col. 26 line 14.) If scale is not set using Step 3.3, then here again mathematically accurate vectors could not be produced and ultimately Solberg is rendered inoperable.

41. It should be noted that Solberg's CAD program would not be able to consume information from the header of a file without substantially changing Solberg's principle operation. In order for Solberg's CAD system to access information from a file header the CAD program would require the computer file be in a format that is native to AUTOCAD, such as a binary proprietary format. There are two acceptable formats: a proprietary format and a supported file exchange format. The proprietary format must be written and read specifically by the CAD application. Solberg does not teach or feasibly suggest modifying CAD applications to support this process. The supported file exchange format is in a .dwg or .dxf format. Solberg does not teach .dwg format and only teaches .dxf as the file name for storing the accurate CAD drawing file. Only the proprietor of the CAD program can change how the CAD system consumes information. Solberg does not teach changing the CAD program to consume information from the header of a file, nor in reality is there any ability/force to change commercially available, proprietary CAD systems.

VIII. EXAMINER FAILS TO PROVIDE AN APPARENT REASON WITH RATIONAL UNDERPINNINGS TO COMBINE SOLBERG WITH RAPPAPORT

42. The office action asserts the motivation to combine Solberg with Rappaport "is to allow users to instantaneously interpret the measurement value (scale information) and allows one to understand or recall with ease the measurement type, measurement location, and etc." (See Rappaport [0074] and [0070].)

43. The office action appears to inaccurately equate Rappaport's term "measurement" to scale. Rappaport uses "measurement" to describe the performance of a network of distributed components (performance metrics) or characteristics of any collection of spatially distributed group of objects. For instance, Rappaport describes measurements with regard to networks such as communication network or a distributed infrastructure network for carrying power, heat, air-conditioning, fluids, and the like or to the physical observation about quality or quantity of objects such as furnishings of a room, quality of paint or inventory of equipment. (Rappaport [0076], [0096] and [0106].) Furthermore, Rappaport does not teach embedding scale. Nowhere in the description of Rappaport's invention is the term "scale" used. Rappaport's single and only use of the term "scale" is made in order to distinguish itself from a prior art reference.[1]

44. Rappaport describes a visual graphical environment where real world objects and environments are represented as approximations with no concern for scale or the implications of securing this data. (Rappaport [0070], [0072], [0078] and [0092].)

45. In contrast, the scale described in the Patent Application is a ratio that in part helps calculate spatial and physical measurements that is taken from the paper drawing. For example, scale is known as inches to inches in a document scale. (Patent Application [0032]-[0044].) I understand the difference between "scale", as claimed in the present Patent Application, and the dimension measurements that the scale helps calculate in the present invention and "measurement" as employed in Rappaport and find the two terms to be different and distinguishable.

---

[1] In an electronic word search conducted on Rappaport, the term "scale" was only used once to distinguish a prior art reference from the Rappaport invention. (See Rappaport [0064]).

46. Also, the language of Rappaport cited as motivation to combine Solberg and Rappaport, when read does not provide an apparent reason with rational underpinning to combine the two references.

47. User's of Rappaport's invention are able to instantly interpret the performance of a network by viewing cues or "graphical display[s] of measurement data" overlaid on the computer representation of the environmental-raster image. (Rappaport [0057] and [0124].) These icons or text strings overlaid on the raster image give context to the value or location of the measurement displayed on the raster image.

48. In contrast, to providing cues to give context, the Patent Application is calculating a true scale measurement using scale information that is embedded in the header of the raster image.

49. As stated above, the combination would not result in a single digital raster image with scale embedded in a dedicated location of the header of the raster file and true scale measurements could not be calculated therefrom as claimed in claims 1, 6, 10 and 15 of the present Patent Application.

50. Also, there is no reason to combine the two references since the combination of the two references would not result in success. As CAD systems do not consume information from the header of any computer file, there would have been any reason to have considered selecting scale information and storing such information in a header of a single digital raster image file. CAD programs store scale information in external library files that can be associated with a CAD file. Furthermore, a large amount of information is typically stored in the CAD library files, thus it would not be feasible to store all of the library file information in the header of a raster file.

51. In addition, if Rappaport were combined with Solberg, the combination of references would likely result in Rappaport's textual strings and/or graphical icons associated with performance metrics being overlaid on a Solberg's intermediate raster file. Solberg would then, if possible, have to recognize and convert these texts and icons as accomplished in Solberg's later steps 4-7. Such a combination again would not result in the invention claimed in

claims 1, 6, 10 and 15 of the Patent Application where a single digital raster image has scale embedded in a dedicated location of the header of the raster file and where true scale measurements are calculated using the embedded scale information.

IX. SECONDARY CONSIDERATIONS

52. Since development of the systems and methods set forth in the present Patent Application, Archaio has achieved substantial commercial success. The inventions have led to multiple sole source contracts in the United States at county, state, and federal levels. In each case, the contracting vehicles point to a unique technology. Additionally, because no bidding process was utilized in any of the cases, Archaio had to demonstrate that no competitive products or offerings were available. Accordingly, I believe that the Archaio inventions are unique, novel and non-obvious.

53. Moreover, as a result of the solution provided by the present Patent Application, a partnership has been established between Archaio and IBM to market products throughout the world. Due to the commercial success and potential market opportunities for various embodiments set forth in the present Patent Application, it is believed that a unique solution has been developed to satisfy the long felt need for systems and methods that efficiently store accurate scale infrastructure plans in an electronic format.

54. During the development of the systems and methods of the present Patent Application, the industry was aware of and familiar with CAD programs and products. However, CAD programs were not incorporated into the invention of the present Patent Application as CAD programs typically utilize their own propriety format making CAD files often not transferable between different systems and programs. (See Exhibits H and I). Another concern with using proprietary formats is the possibility of the owner of that particular format going out of business thus leaving the format unsupported in the future. (See Exhibit I.) The Patent Application overcomes these concerns as it utilizes non-proprietary files with scale information embedded in a header. These non-proprietary raster files are easily transferable between different systems and software programs and readily available for use by a wide variety of different individuals. The files can be opened and the embedded scale information may be quickly accessed following scanning in order to determine true scale measurements. Also, CAD

programs were not incorporated into the invention of the present Patent Application as CAD programs store scale information in external library files thereby increasing the likelihood of information loss or disassociation. Accordingly, CAD programs were not considered when developing the invention of the Present Application.

55. The Patent Application provides a single phase validation where, by embedding scale of the initial document in the header, measurements can be calculated by directly using the embedded scale information stored in the dedicated location of the header of a file. Thus, measurement results can be validated simply by confirming the accuracy of the scale embedded in the header. In contrast, the prior art inventions require a minimum 3-phase validation. For instance, in the Solberg invention validation can only be performed by a multi-phase validation of the Raster to Vector CAD conversion process. Solberg's multi-phase validation includes storing scale in the CAD file, manually inputting the scale information within the CAD system and finally reviewing the source document. Furthermore, validation of the Solberg method is impossible to perform if the file has been shared and the party does not have access to the source document.

56. The Patent Application provides a secure and efficient means of ensuring accuracy because measurements are calculated by directly using the scale embedded in the file, specifically in the dedicated location of a header of the file. Thus the Present Application ensures the calculation of accurate measurements even after file sharing and distribution.

57. The Patent Application provides an immediately intelligible raster file from which true scale measurement can be calculated.

58. The U.S. Army has allocated $34 Million to CACI International, Inc. as a prime federal systems integrator to provide three-dimensional renderings of blueprints and architectural drawings that enhance the Army's ability to ensure comprehensive facilities support and protection. CACI has not allocated a vendor to the project and are in partnership discussions with Archaio as the present Patent Application and related applications are the only existing solution to quickly, efficiently and accurately produce rendering of blueprints and architectural drawings, viewable precisely to scale.

59. The present Patent Application provides a non-proprietary solution that provides several unique benefits over the proprietary CAD systems such as those non-limiting examples of benefits cited above. The resultant non-proprietary file can be distributed and utilized by collaborating agencies, departments, and companies without the need to purchase a CAD system or invest in training and maintaining CAD users and environment.

60. The ability to share and corroborate accurate information quickly is vital to our National Security and Emergency and Disaster Management Centers. There is an urgent and immediate need to quickly and efficiently prepare and respond based on a Common Operating Pictures (COP). The unique capabilities provided by the Patent Application delivers a comprehensive internal situation awareness of the COP which can be used to save lives and protect property.

X. <u>ATTESTATION</u>

61. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with knowledge that willful false statements and the like so made are punishable by fine or imprisonment or both under Section 1001 of Title 18 of the United States Code and that willful false statements may jeopardize the validity of the application or any patent issued thereon.

Date: August 17, 2009

Lawrence A. White

Lawrence A. White

EXHIBIT

A

Lawrence A White

(404)424-3084

2750 Towne Village Dr, Duluth, GA 30097

lwhite@archaio.com

Professional Profile

Over 25 years of experience applying Information Technology to solve core business problems and processes. Highly skilled Solution Architect with significant experience with CAD development, CAD integration, Emergency Management, and Services Oriented Architectures (SOA).

Consulting Experience:	Processing Re-Engineering, Organizational Change Management
Technical Skills:	System Architecture, System Integration, High Availability, Disaster Management
Software Engineering:	C/C++, Object-Oriented Programming, Pascal, Fortran, Assembler
CAD Software:	CEAL, AutoCAD, MicroStation, Intergraph, RDS, CADAM, CATIA

Professional Accomplishments

Gauteng Provincial Disaster Management and Emergency Operations Centre

Lead consultant on the design, architecture, implementation and operational readiness of the Gauteng Provincial Disaster Management and Emergency Operations Centre in Midrange, South Africa, serving Pretoria, Johannesburg and the greater Gauteng Province. The Centre's focus is on disaster risk reduction and effective emergency preparedness of emergency services from all sectors. Including pre-planning of infrastructural developments, assessing capacity and identifying gaps of line functions to respond to major incidents. The Centre will provide real-time monitoring of critical facility and will also be used for the 2010 World Cup. The Centre provides effective response to major incidents and disasters that entail the development of effective response plans, ensuring the implementation of line function response plans, testing and training through desktop exercises of response plans, as well as, the recruitment and training of emergency service volunteers.

St Louis Area Regional Response System

Lead consultant on the design, architecture, implementation, operational readiness and deployment of the St Louis Area Regional Response System (STARRS). The STARRS system provides emergency response, collaboration and preparedness across eight counties and two states within DHS UASI region centric to St Louis. STARRS is design to support preparedness and response to critical incidents such as, natural disasters, pandemics, accidents and intentional acts like chemical, biological, radiological, nuclear or explosive (CBRNE) events.

New Jersey State Police

Lead consultant on the design, architecture, and implementation of the State of New Jersey Emergency Prepardness Information Network (EPINet) OneStop. The New Jersey EPINet OneStop is a secure and personalized gateway providing simplified, integrated access to people, information, applications, and business processes for those involved in law enforcement, homeland security, emergency preparedness and response. The New Jersey EPINet OneStop is part of the EPINet program. EPINet is New Jersey's program for Implementing Information Technology solutions that span the Homeland Security Law Enforcement, Emergency Management and First Response communities of interest. The features include access to EPINet subsystems including:

- E-Team - the New Jersey Statewide incident management system that provides situational awareness
- RDDB - a database used to identify and catalog all the resources available within the State of New Jersey
- EMV - a powerful tool for querying and visualizing spatial information
- Searchable information clearinghouse for data, contracts, services, applications, and resources
- Multiple profiles
- Live chat/meeting capabilities
- Support for multiple clients (desktop, web, mobile)

IBM Corporation

Engineering/Scientific Specialist supporting Desert Research Institute, Department of Defense Engineering Research and all major Engineering Research companies in the Western United States. Extensive CAD deployment, integration and innovation. Experience with CADAM, CATIA, MicroStation, AutoCad, and CEAL.

US Federal

IBM Executive responsibility expanded the IBM Federal Systems Division portfolio to include unique and innovative IBM technology services.

Puerto Rico Highway Authority

Lead consultant on the design, architecture, integration and implementation of a comprehensive engineering system for the Puerto Rico Highway Authority (PRHA). The project included all engineering department of the PRHA: Survey, Photogrammetry, Road & Highway Design, and Bridge Design. The core CAD components of the project were comprised of CEAL, Intergraph, CADAM, CATIA, and GTSUDL.

Top 5 Engineering Firm and 17 Statewide Department of Transporation Agencies

Lead engineer supporting the top 5 national engineering firms in the US and 17 US Statewide Department of Transportations (DOT). Responsible for development, deployment and integration of lead CAD innovations for the Civil Engineering Community.

Major Awards and Certifications

IBM Blue Ribbon Award
CEO award presented by Lou Gerstner for significant contribution to the development of the IBM Global Services blueprint and revitalization of IBM.

IBM Magic of Leadership Award
General Manager award presented by Ginny Rometti to the top Global Services Account Executive Team that serves as a model of teamwork for IBM's top 50 integrated accounts.

IBM Golden Circle Awards (multiple)
Presented to the top 1% performance for revenue, profit and growth contributions.

IBM System Engineer Excellence Awards (multiple)
Presented to the top 1% performance for technical excellence.

IBM Certified Senior IT Professional
In 1995 Became the first IBM Certified Senior IT Professional. Was appointed a by Senior Management a member of the IBM IT Professional Certified Board.

IBM Rookie of the Year
IBM Sr. Vice President Award presented by Ned Lautenbach. Presented to the top first year contributor of approximately 20,000 professional hires and 30,000 college hires.

HEEP (Highway Engineer Exchange Program)
Multiple guest and keynote speaker engagements

AASHTO (American Association of State Highway Transportation Officials)
Multiple guest and keynote speaker engagements

India Minister of Science & Technology Security Round Table
One of 18 participates in a day long roundtable session on the role of I/T in National Security. Hosted in Dehli, India by the India Minister of Science & Technology.

Work History

CTO	Archaio, L.L.C.	01/2009 - Present
VP Enterprise Planning	VirtualAgility	04/2007- 01/2009
Director, Public Safety Solutions	Paaridian Technologies	10/2004 - 04/2007
Principle	IBM Corporation	11/1992 - 09/2004
Engineering/Scientific SE	IBM Corporation	09/1989 - 10/1992
VP Operations	CLM System, Inc.	08/1989 - 09/1985

Education

Computer Science	University of South FL	09/1980 - 12/1985
Management Information Systems	California Coast University	01/1990 - 12/1991

EXHIBIT

B

Evaluation Report

OFFICE OF THE INSPECTOR GENERAL

EVALUATION OF AUTOMATED DOCUMENT CONVERSION IMPLEMENTATION

Report No. 96-153

June 10, 1996

DEPARTMENT OF DEFENSE

Additional Copies

To obtain additional copies of this report, contact the Secondary Reports Distribution Unit of the Analysis, Planning, and Technical Support Directorate at (703) 604-8937 (DSN 664-8937) or FAX (703) 604-8932.

Suggestions for Future Audits

To suggest ideas for or to request future audits or evaluations, contact the Planning and Coordination Branch of the Analysis, Planning, and Technical Support Directorate at (703) 604-8939 (DSN 664-8939) or FAX (703) 604-8932. Ideas and requests can also be mailed to:

OAIG-AUD (ATTN: APTS Audit Suggestions)
Inspector General, Department of Defense
400 Army Navy Drive (Room 801)
Arlington, Virginia 22202-2884

Defense Hotline

To report fraud, waste, or abuse, contact the Defense Hotline by calling (800) 424-9098; by sending an electronic message to Hotline@DODIG.OSD.MIL; or by writing the Defense Hotline, The Pentagon, Washington, D.C. 20301-1900. The identity of each writer and caller is fully protected.

Acronyms

ADCS	Automated Document Conversion System
CAD	Computer-Aided Design
CALS	Continuous Acquisition and Life-Cycle Support
DLA	Defense Logistics Agency
DPS	Defense Printing Service
DUSD(L)	Deputy Under Secretary of Defense (Logistics)
DXF	Data Exchange Format
IGES	Initial Graphics Exchange Specification
JEDMICS	Joint Engineering Drawing Management and Information Control System
OASD(C^3I)	Office of Assistant Secretary of Defense (Command, Control Communications, and Intelligence)
PC	Personal Computer

INSPECTOR GENERAL
DEPARTMENT OF DEFENSE
400 ARMY NAVY DRIVE
ARLINGTON, VIRGINIA 22202-2884

Report No. 96-153 June 10, 1996

MEMORANDUM FOR UNDER SECRETARY OF DEFENSE FOR ACQUISITION AND TECHNOLOGY
ASSISTANT SECRETARY OF DEFENSE FOR COMMAND, CONTROL, COMMUNICATIONS AND INTELLIGENCE
DIRECTOR, DEFENSE LOGISTICS AGENCY
DIRECTOR, DEFENSE INFORMATION SYSTEMS AGENCY

SUBJECT: Evaluation of Automated Document Conversion Implementation (Project No. 6PT-5003)

Introduction

We are providing this report for information and use. By agreement with the Principal Deputy Under Secretary of Defense for Acquisition and Technology, we evaluated the concerns raised by Congressman Hunter, Chairman of the Subcommittee on Military Procurement, House Committee on National Security, and a contractor, AUDRE, Inc., regarding an Automated Document Conversion System (ADCS) for engineering drawings. AUDRE, Inc., was one of the vendors whose software the Defense Printing Service (DPS) assessed as a candidate for the DoD ADCS. The Chief Executive Officer of AUDRE, Inc. told Congressman Hunter that the Defense Logistics Agency (DLA) was not buying the AUDRE automated document conversion system even though DPS assessed it as the best candidate system. In the National Defense Authorization Act for Fiscal Year 1996, the conferees expressed their concerns that DoD was not making progress to achieve "major cost savings" through adopting the automated document conversion technology. Congressman Hunter specifically questioned why the DoD was not using a previous $20 million appropriation to automate the conversion of engineering drawings.

Evaluation Results

We determined that:

o Limited demand exists for conversion of legacy hard copy engineering drawings to vector format.

o The state-of-the-art in automated document conversion technology has not progressed to the level that allows an agency to convert a rasterized drawing into its vector equivalent without human intervention.

o The DoD has developed sound policies and an effective strategy through which to implement automated document conversion in a cost-effective manner.

o The DoD has prudently expended document conversion funds.

Evaluation Objective

Our objective was to evaluate the degree to which DoD has implemented automated document conversion of engineering drawings. Therefore, we examined:

o whether demand exists for automated document conversion of engineering drawings,

o the state-of-the-art in automated document conversion technology and whether the technology is cost-effective,

o whether DoD has established automated document conversion policy that will ensure the cost-effective conversion of engineering drawings, and

o whether DoD has prudently applied automated document conversion funding.

Scope and Methodology

The scope of our evaluation included a review of document conversion within the Continuous Acquisition and Life-Cycle Support (CALS) Joint Engineering Drawing Management and Information Control System (JEDMICS) Program Management Office and at field agencies. We interviewed agency officials and observed agencies' document conversion activities. We also reviewed the Automated Document Conversion Master Plan published by the Office of the Assistant Secretary of Defense (Command, Control, Communications, and Intelligence) (OASD[C^3I]). We began the evaluation October 23, 1995, and completed it December 21, 1995.

We attended the CALS International Expo 95. We visited the Office of the Deputy Under Secretary of Defense (Logistics) (DUSD[L]), who recommended we contact specific organizations that are involved first-hand in document conversion. We met with representatives from the JEDMICS Program Management Office; AUDRE, Inc.; the Oklahoma City Air Logistics Center; DPS at Port Hueneme, California; and the Naval Undersea Warfare Center at Keyport, Washington. We also met with the ADCS action officer from the Headquarters DPS Plans, Policy, and Technology Assessment Office (Enclosure 2).

Three basic legacy document types within DoD are subject to conversion: technical publications, maps, and engineering drawings. Because the congressional concerns focus on the conversion of engineering drawings, we only evaluated the DoD efforts to convert legacy engineering drawings.

Prior Audits and Other Reviews

Since June 1994, three Inspector General Reports have related to automated document conversion of data into digitized format. Enclosure 1 discusses the prior reports.

Background

DoD has millions of legacy weapon systems drawings in hard copy format. DoD uses some of these documents in weapon systems' upgrades and in the development of new weapon systems. To edit drawings for upgraded or new weapon systems, agencies must be able to use the drawings in computer-aided design (CAD) systems. Before agencies can use these documents in their CAD systems, they must convert the documents to a digital format that a CAD system can edit. To convert these documents, engineers may have to spend numerous hours to manually trace a scanned image of a drawing or completely reconstruct a drawing. Congress believes that DoD is currently using thousands of workstations to convert legacy documents in this manner and that DoD is spending hundreds of millions of dollars a year to convert these documents.

In FY 1994, Congress appropriated $14 million to the Defense Logistics Agency to competitively procure an ADCS. Congress believed that after scanning a document, an ADCS would eliminate the need for further human intervention in the conversion process. The ADCS could cut conversion costs by reducing the labor needed to convert documents. Therefore, agencies would be able to efficiently convert hard copy drawings or drawings on aperture cards (punched cards on which a microfilmed document is mounted) into intelligent digital files using automated rather than manual methods. The data file output by the ADCS was to follow the standards of the CALS initiative.

CALS is the DoD and industry technological initiative to integrate and use automated digital technical data for weapon systems acquisition, design, manufacture, and support. The objective of CALS is to facilitate the transition from the current paper-intensive weapon systems acquisition to an environment that provides for the generation, exchange, management, and use of digital data. An ADCS would provide a more cost-effective conversion of hard copy engineering drawings into this intelligent digital technical data instead of conversion through manual processes. DoD could then use this intelligent digital data to reduce weapon system acquisition times and costs. DoD has adopted a two-stage process to convert hard copy drawings to an intelligent digital format.

The first stage of the conversion process is to scan the drawing to create a digital raster file of the drawing. A raster image is a bit-mapped representation of a drawing: a digital photograph. The degree of resolution is measured in dots per inch. Raster images can only be edited by adjusting the values of individual dots; therefore, they do not provide an editable CAD-ready file. They also require human interpretation as do hard copy drawings and drawings on aperture cards. However, raster files are good for archiving and for print-on-demand requirements.

The second stage of the conversion process converts the raster file to a vector format. A vector file is the presentation or storage of images as sequences of line segments. These data files consist of geometrically accurate and precise representations of the product, together with associated annotations such as dimensions and tolerances. Vector files can be scaled, which means one can zoom in on the details of a drawing. Also, they are more easily edited than raster files. Creating a vector file requires converting the lines and arcs in a scanned bitmap to the equivalent structure in a CAD system. Before automated document conversion systems, agencies were only able to do this conversion by manually tracing or reconstructing the drawing. The intent of the ADCS initiative was to offer a system to automatically (without human intervention) convert a scanned raster image to a vector format.

Toward that objective, DLA sponsored a state-of-the-art assessment through the DPS in April 1994. DPS conducted this assessment to identify candidate vendor automated document conversion technology packages that convert technical publications, maps, and engineering drawings. The assessment included six vendors who offered products able to automatically convert engineering drawings to vector format. Of the six vendor products, DPS assessed the AUDRE Automatic Conversion System as the best candidate engineering drawing ADCS. DPS would later test the AUDRE Automatic Conversion System for Congress.

In FY 1995, Congress appropriated $20 million to DoD to implement the ADCS for engineering drawings Defense-wide. Congress also directed that the OASD(C^3I) establish and implement a master plan for all acquisitions of automated document conversion systems, equipment, and technologies.

Between July and November 1994, DPS tested the AUDRE Automatic Conversion System. The test showed automation-assisted labor savings could result from using the AUDRE Automatic Conversion System to convert engineering drawings to vector format. However, DPS added that automated document conversion was not mature enough to completely replace trained engineers with production operators. After DPS tested the AUDRE Automatic Conversion System, it stored the tested AUDRE software packages at various DPS locations. However, these systems were not added to the DPS inventory for use in document conversion because of a lack of user requirements according to DPS.

Discussion

Demand for Automated Document Conversion. Our evaluation of automated document conversion began with determining whether the universe of documents eligible for document conversion contains any demand for vectorization of legacy hard copy engineering drawings.

DoD estimates that as many as 100 million engineering drawings exist. The universe of engineering drawings includes architectural engineering, electrical engineering, and mechanical engineering drawings. OASD(C^3I) officials stated that no credible estimate of organization records eligible for conversion exists.

However, DoD estimates that only 10 percent of raster engineering drawings require conversion to vector format. See Figure 1. OASD(C^3I) officials indicate that the many technology-related marketing claims may have generated an artificial demand for conversion of DoD paper records to digital media formats.

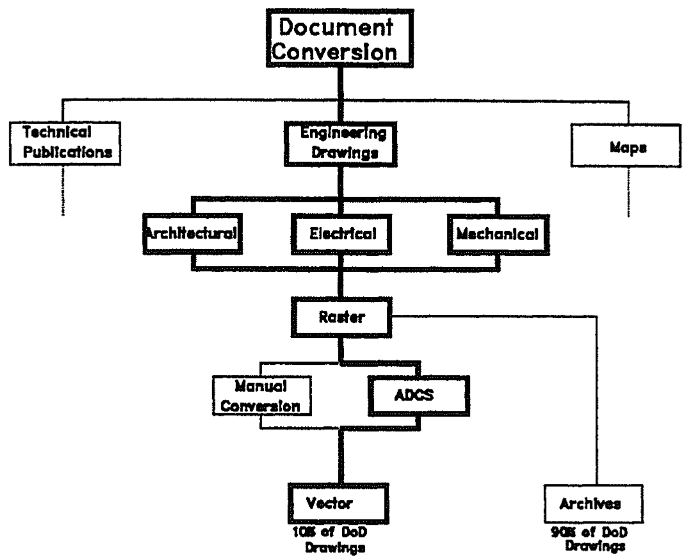


Figure 1. DoD Estimate of Engineering Drawings Requiring Conversion to Vector Format

None of the agencies we visited expressed a requirement to bulk convert documents to vector format; they implemented conversion requirements as needed. One agency even found bulk conversion of legacy hard copy documents to digital *raster* format to be unnecessary because it no longer uses many of the drawings. Therefore, agencies must first determine which of their legacy drawings in general they might use, if any, before they consider which drawings they might need to convert to a vector format.

According to OASD(C^3I), deciding which documents to convert is crucial to any conversion project. Military mission and business requirements and a business case that clearly explains the functional and economic benefits anticipated from conversion will guide the conversion decisions. In addition to military mission and business requirements, these decisions include the timing of automated conversion and the justification for automated conversion. The decisions are delegated to responsible functional and DoD Component officials.

Conclusion. Limited demand exists for conversion of legacy hard copy engineering drawings to vector format.

Cost-effectiveness of ADCS Technology. To evaluate ADCS technology in the conversion of engineering drawings from raster to vector format, we examined current available conversion technology and the future of automated document conversion technology. We also looked at the minimum operator skill and knowledge required to provide a cost-effective conversion, the operating system and hardware platform requirements that will most cost-effectively use current DoD platforms within the engineering community, and which data format standards the ADCS should produce to provide a cost-effective digital file for use in the acquisition life-cycle.

Current Conversion Technology. To convert a hard copy document to a vector file, DoD has adopted a two-stage process. The first stage of conversion establishes an interoperable baseline digital raster format, allowing agencies to share information electronically. The second stage further processes the digitized image into a more complex digital vector format if required by the target application. This flexible two-stage approach extends the potential for reuse of a converted document to satisfy different user requirements for the same document. Within this two-stage conversion process, we identified several steps necessary to achieve a quality CAD-ready vector file that accurately depicts the original drawing.

The first step is to scan the hard copy drawing or aperture card to create a digital raster image. The second step involves a quality assurance function to ensure the quality of the scanned raster image. This process may involve deskewing and despeckling the image. The third step is to convert the raster file to a vector file. This step identifies and captures the different parts of the image. This step in the conversion process is automated, which produces an initial vector image that the ADCS operator can then edit. In the fourth step, the ADCS operator performs a quality assurance check and edit of the initial vector image. The ADCS operator would then pass the vector file to the using engineer for a final quality assurance check to ensure that all parts of the image were captured and converted correctly. This step may involve editing objects or symbols that the ADCS may have misrecognized in the third step and the ADCS operator missed or was unable to interpret in the fourth step. Since this fifth step should be done by a subject matter expert, such as an engineer, it could most cost-effectively be accomplished in the using agency's target application that the engineer is familiar with, such as AUTOCAD. See Figure 2.

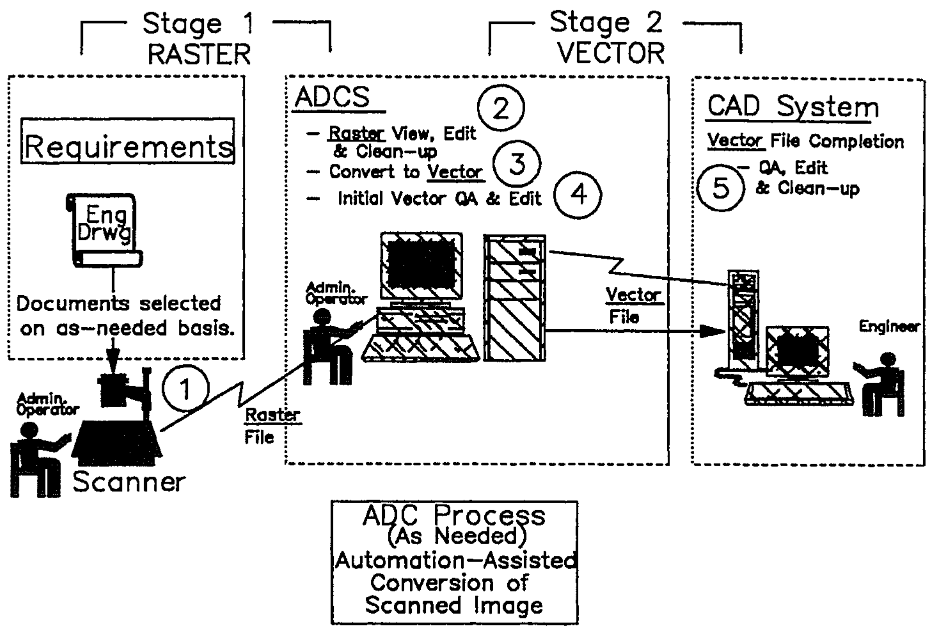


Figure 2. Two-stage, Five-Step Automated Document Conversion Process

The condition of the source document and the type of drawing determine the amount of human interaction by the ADCS operator and the subject matter expert. Because the second, fourth, and fifth steps require the interaction of ADCS operators and subject matter experts, such as engineers, to generate an acceptable CAD-ready file, an ADCS is considered to provide only an automation-assisted conversion capability rather than a fully automated conversion capability. However, one of the agencies we talked with that uses production operators (as opposed to engineers) stated that prior to the AUDRE Automatic Conversion System, it would not convert engineering drawings. The agency said that the AUDRE system is its only option for CAD-ready conversions. The only current alternative to automation-assisted drawing conversion is tracing or completely reconstructing a drawing. Another agency, which was not familiar with the AUDRE Automatic Conversion System, continues to use manual conversion methods.

AUDRE executives explained that the sole purpose of its ADCS is to be a conduit to other applications, such as AUTOCAD, so that engineers can complete the conversion and get on with their upgrades and changes more quickly. An ADCS cuts the conversion time so that engineers can access an editable file in a more timely manner so that they could then change and upgrade a drawing in their target CAD environment. One agency plans to use the AUDRE Automatic Conversion System to provide one of its customers a Data Exchange Format (DXF) file to convert to another format using another application. Another agency has been using AUDRE for 2 to 3 years as a

routine part of its production operations. The agency provides its customers with an editable vector file so that the customer engineers will not have to redraw the drawing from scratch. The agency explained that a conversion process that would take an engineer weeks to complete using tracing methods would only take it a day using the AUDRE Automatic Conversion System.

Future Conversion Technology. According to experts in electronic document conversion technology, fully automated document conversion systems are not likely to be developed soon. Therefore, the engineering industry is likely to be using automation-assisted conversion technology for some time.

ADCS Operator Skill and Knowledge. The ADCS Test found that blue collar administrative personnel with basic computer skills and no engineering drawing experience can attain the skill necessary to cost-effectively convert engineering drawings. Compared to the manual alternatives, the ADCS Test Report indicated that the most productive administrative ADCS operators achieved better throughput times and quality than either the professional engineers or the commercial conversion operators using manual methods. The report added that costs to convert a typical drawing vary from $200 if redrawn by engineers to $119 if converted through existing commercial sources to $85 if converted by experienced operators using automation-assisted conversion techniques. The report estimates that labor savings range from 20 percent to 50 percent over the manual redraw methods that professional engineers use.

One agency has been using the AUDRE Automatic Conversion System for the last 2 to 3 years. Most of its operators are high school graduates with a printing operation background rather than a computer or data entry background. The operators learned the system in about 3 weeks. The agency stated that its operators have converted in a day what their customer agency engineers would take weeks to convert using manual tracing methods.

ADCS Operating System and Hardware Platform Requirements. The AUDRE Automatic Conversion System can operate on a Personal Computer (PC) (Sun Solaris operating system) or UNIX (Hewlett Packard or Sun) platform. The hardware platform is not a factor in producing a CALS-standard vector output. In addition, the ADCS can output the vector file in the appropriate format for the target application. However, AUDRE executives stated that the PC platform presents memory management problems. They added that their software can co-exist on a system already containing another application such as AUTOCAD.

ADCS Digital Vector Data Format Standards. The digital data produced by the ADCS was required to follow the CALS digital data format standards. The Initial Graphics Exchange Specification (IGES) is the CALS digital vector format standard for CAD system engineering drawings.

The IGES file format treats the data required to describe and communicate the essential characteristics of physical objects as a file of entities. Each entity is represented in an application-independent format, which can be mapped to and from a native representation of a specific CAD system. Because IGES is

available in five versions, problems can result when trying to transfer IGES files between different CAD systems. The transfer can result in mis-mappings and the loss of data.

OASD(C^3I) officials state that DoD recognizes de facto standards until formal standards are developed and adopted by recognized standards organizations. The de facto industry CAD standard format is DXF. DPS chose to use the DXF format during the ADCS test instead of the IGES format because it is the most acceptable common CAD file format within the engineering industry.

The AUDRE Automatic Conversion System can readily output to either format. However, the two agencies we visited that use the AUDRE Automatic Conversion System produce digital data in the DXF format instead of the IGES format. These agencies use the DXF format because it is a neutral file format, which the end-user can convert to any other format.

Conclusion. The Automated Document Conversion System provides a cost-effective automation-assisted conversion capability using administrative operators. Therefore, the ADCS can provide significant labor savings over highly skilled engineers manually tracing or reconstructing drawings. The potential for savings depends partly on whether an agency has a sufficient demand to convert documents to vector format (See "Demand for Automated Document Conversion," page 8).

Automated Document Conversion Policy. We reviewed whether DoD has established an automated document conversion policy that will ensure the cost-effective conversion of engineering drawings.

In the FY 1995 Defense Appropriation Act, Congress directed the OASD(C^3I) to establish and implement a master plan for all acquisitions of automated document conversion systems, equipment, and technologies. In April 1995, OASD(C^3I) published the Automated Document Conversion Master Plan (the Master Plan).

The Master Plan provides strategic guidance for all automated document conversion acquisitions within DoD. It focuses on conversion from paper or microform to digital formats. The Master Plan addresses three main areas:

o the "DoD Conversion Environment," which summarizes the mission and business needs for automated document conversion;

o the "DoD Conversion Strategy," which describes the DoD strategy for achieving a consistent approach to automated document conversion; and

o "DoD Roles and Responsibilities" applicable to automated document conversion.

One theme of the "DoD Conversion Environment" is for agencies to "follow existing policy." The Master Plan views document conversion as an activity within the "records management" business process. It also states that the requirement or business need for document conversion must be justified using existing Corporate Information Management principles and automated information system life-cycle management policy.

One component of the "DoD Conversion Strategy" focuses on the management of automated document conversion system acquisitions and requirements. This management component provides agency guidance to determine whether a proposed automated document conversion acquisition meets operational requirements and produces sufficient cost savings. This guidance provides agencies with decision criteria in four areas: Requirements Determination, Cost Justification, Document Candidate Selection, and Technical Capability. Agencies should consider these criteria before deciding to proceed with automated document conversion.

The Office of the DUSD(L) has also taken steps to ensure the cost-effective conversion of engineering drawings. The Office of the DUSD(L) issued a data call in August 1995 to the Services regarding what their experiences have been with document conversion, their data conversion requirements, their experience with the cost-effectiveness of document conversion, and what software and hardware they use in document conversion. The Office of the DUSD(L) will also decide whether procuring and fielding a PC-based ADCS is worthwhile given the preponderance of UNIX workstations in the DoD engineering community.

Congress has requested that the Office of the DUSD(L) provide it the results of these efforts. Specifically, the report will address:

o the logistics community's requirements and strategy for raster to vector conversion,

o the drawing document universe in the field,

o the number of engineers in the field,

o the level of existing UNIX and PC platforms, and

o an acquisition plan for the software and hardware and who the vendors will be for each.

Conclusion. The DoD has developed sound policies and an effective strategy through which to implement automated document conversion in a cost-effective manner.

Application of Automated Document Conversion Funding. We reviewed whether DoD has applied automated document conversion funding in a prudent manner.

With the $14 million appropriated to DLA in 1994, DLA procured ADCS hardware and software and met other costs, such as salaries, training, travel, and conversion, associated with its evaluation of the ADCS technology. Also, DLA funded an independent operational appraisal of the ADCS technology by the Defense Information Systems Agency Joint Interoperability Test Command.

Of the $20 million FY 1995 appropriation, DLA allocated $10 million for procurement and $10 million for operation and maintenance. As of December 8, 1995, DLA had spent $7.5 million of the procurement allocation to procure ADCS hardware and software. In May 1995, DLA procured 50 AUDRE software packages and 34 Hewlett Packard workstations for some field agencies to use in their document conversion efforts. In the first quarter of FY 1996, in response to requests from Congress to evaluate raster to vector conversion products in the field, DLA procured an additional 20 AUDRE workstation-based systems, 100 AUDRE PC-based systems, and 100 each PC-based systems of four other vendors for the Services to evaluate.

Also, DLA provided $2.2 million of the Operation & Maintenance allocation to one specific agency to convert its documents from raster to vector. DLA had requested the Services determine whether or not they had requirements for raster to vector conversion; only one Service responded. As a result, DLA wants to spend the remaining $7.9 million on raster to vector conversion within a single weapon system program.

Conclusion. DoD has prudently expended document conversion funds.

Management Comments

We provided a draft of this report to you on April 25, 1996. Because this report contains no findings or recommendations, written comments were not required and none were received. Therefore, we are publishing this memorandum report in final form.

We appreciate the courtesies extended to the evaluation staff. If you have questions on this report, please contact Mr. Kenneth H. Stavenjord, Technical Director, at (703) 604-8952 (DSN 664-8952) or Mr. Gregory R. Donnellon, Evaluation Project Manager, at (703) 604-8946 (DSN 664-8946). See Enclosure 3 for the report distribution. The evaluation team members are listed inside the back cover.

Robert J. Lieberman

Robert J. Lieberman
Assistant Inspector General
for Auditing

Enclosure

Summary of Prior Audits and Other Reviews

Three Inspector General, DoD reports covered issues related to this evaluation.

Report No. 95-060, "Digital Mapping, Charting and Geodesy Data Standardization," December 19, 1994, found that the Defense Mapping Agency had taken positive actions to standardize digital mapping, charting, and geodesy data. The purpose of standardization was to promote electronic transfer between military systems and to promote system compatibility and interoperability. The report made no recommendations for corrective actions and no comments were made in response to the final report.

Report No. 95-043, "Management of the Digital Production System Development at the Defense Mapping Agency," November 28, 1994, found that the Agency did not identify customer requirements, did not analyze the cause of software problems, and did not correct configuration management deficiencies. The report recommended that the Defense Mapping Agency improve its product specification development and revisit its problem reporting. The report also recommended corrective actions on configuration management procedures and the Digital Production System. Finally, the report recommended a Milestone IV (Major Modification Approval) review of the Digital Production System. Management concurred with the recommendations regarding program management and agreed to review the Digital Production System.

Report No. 94-INS-05, "Management of Digitized Technical Data," July 8, 1994, found a lack of management and clear and consistent guidance from DoD. The report specifically criticized management of the CALS initiative because CALS was not defined as a strategy or as a program. The lack of definition created an ineffective management structure, late allocation of funds, a lack of policies on reimbursement for operating funds, and a lack of specific guidelines needed to acquire and manage digitized technical data. The Joint CALS system had similar problems. The report recommended several changes, including greater structuring of CALS and Joint CALS management, writing action plans for implementation, changes to regulations, greater oversight, and changes to data standards. DoD agreed to most recommended changes; however, the DoD rejected the recommendation to modify the funding of CALS by removing the program from the Defense Business Operations Fund and moving it under direct appropriations.

Organizations Visited or Contacted

Office of the Secretary of Defense

Office of the Deputy Under Secretary of Defense (Logistics), Washington, DC

Department of the Navy

Naval Undersea Warfare Center, Keyport, WA
Defense Printing Service, Washington, DC

Department of the Air Force

Oklahoma City Air Logistics Center, Tinker Air Force Base, OK

Other Defense Organizations

Defense Printing Service, Port Hueneme, CA
Program Management Office, Joint Engineering Drawing Management and Information Control System, Arlington, VA

Non-Government Organization

AUDRE, Incorporated, San Diego, CA

Report Distribution

Office of the Secretary of Defense

Under Secretary of Defense for Acquisition and Technology
Director, Defense Logistics Studies Information Exchange
Under Secretary of Defense (Comptroller)
Deputy Chief Financial Officer
Deputy Comptroller (Program/Budget)
Assistant Secretary of Defense (Command, Control, Communications and Intelligence)
Deputy Under Secretary of Defense (Logistics)
Director, Continuous Acquisition and Life-Cycle Support Office
Assistant to the Secretary of Defense (Public Affairs)

Department of the Army

Auditor General, Department of the Army

Department of the Navy

Assistant Secretary of the Navy (Financial Management and Comptroller)
Director, Defense Printing Service
Auditor General, Department of the Navy
Commander, Naval Undersea Warfare Center, Division Keyport

Department of the Air Force

Assistant Secretary of the Air Force (Financial Management and Comptroller)
Auditor General, Department of the Air Force
Commander, Oklahoma City Air Logistics Center

Other Defense Organizations

Director, Defense Contract Audit Agency
Director, Defense Information Systems Agency
Director, Defense Logistics Agency
Director, National Security Agency
Inspector General, National Security Agency
Inspector General, Defense Intelligence Agency
Program Manager, Joint Engineering Drawing Management and Information Control System

Non-Defense Federal Organizations and Individuals

Office of Management and Budget
Technical Information Center, National Security and International Affairs Division, General Accounting Office

Chairman and ranking minority member of each of the following congressional committees and subcommittees:

- Senate Committee on Appropriations
- Senate Subcommittee on Defense, Committee on Appropriations
- Senate Committee on Armed Services
- Senate Committee on Governmental Affairs
- House Committee on Appropriations
- House Subcommittee on National Security, Committee on Appropriations
- House Committee on Government Reform and Oversight
- House Subcommittee on National Security, International Affairs, and Criminal Justice, Committee on Government Reform and Oversight
- House Committee on National Security
- House Subcommittee on Military Procurement, Committee on National Security

Evaluation Team Members

This report was prepared by the Analysis, Planning and Technical Support Directorate, Office of the Assistant Inspector General for Auditing, DoD.

Michael G. Huston
Kenneth H. Stavenjord
Gregory R. Donnellon
W. Marvin Stith
Danny B. Convis

EXHIBIT

C

TracTrix FAQs on converting from scans to CAD vectors and other raster image issues.

What's the difference between raster and vector files?

A raster image is roughly equivalent to a newspaper photograph. If you look closely you see that it is made up of lots of small dots. The density of these dots is referred to by resolution, measured in dots per inch (stated as 'dpi'). Raster files are stored in many file formats. TIFF is a common example.

Vectors are used to describe specific objects defining their size, position and geometry. The image seen on screen in a CAD application is only a visual representation of the vector file. The vector information in a CAD file is mathematical. It is not stored as an image.

So, for example, a vector line is an absolute line connecting two or more points; a vector curve is described by source coordinates, a polynomial equation and end point coordinates.

Vector information is stored in many different file formats, one of which is DXF. The primary function of the DXF file is to permit CAD vector information to be exchanged between different CAD applications.

Do scans work in a CAD program?

Scanners produce raster images, not the vectors used in CAD programs. When CAD software allows import of a raster, it will display it as a raster (i.e. dot based) image within the CAD document.

The lines in a raster image in a CAD document cannot be manipulated in the CAD program. The CAD program will not convert it to vectors unless assisted by a conversion program like TracTrix.

How can I use my desktop scanner to create CAD files?

Paper to CAD (raster to vector) software such as TracTrix take your scans, convert them to vectors and output them as a DWG, DXF, IGES, HPGL or other vector file format. In the case of TracTrix this can be done within AutoCAD itself or outside of AutoCAD.

What's the difference between pixels and dots per inch? Is a dot a pixel?

A pixel is a minute circular point of light about 10 or 12 thousandths of an inch in diameter (sometimes stated as 10 or 12 'mils') created on a monitor screen.

A dot is a sample taken by a scanner for representation in a raster image. A 300 dots per inch (dpi) scan takes snapshots of dots which are about 3 mils across (3,33333....mils to be exact). To represent a 300 dpi scan on screen at a 1:1 scale your computer typically has to consolidate groups of 16 dots (each 3.33... mils diameter) into a single pixel (10 - 12 mils diameter)

This means that the raster image you see on screen is only an approximation of the underlying digital 'dpi' information from your scan. And your scan itself is an approximation. Sometimes the screen appearance can be disconcerting.

Why can't I just edit my raster files and not have to bother taking them to vector?

Raster files are excellent for visual purposes, for Web posting and distribution, for archiving and for security in release control (see Trix RasterServer) - after all they are a picture of the original image just like a print. TracTrix does provide raster

editing tools so you can edit and clean up your scans. But raster editing can be tedious if there are considerable modifications to be done - especially if you are used to the power of a CAD package.

And all CAD software uses vector objects instead of rasters. The primary benefits are ability to precisely describe, create, scale and manipulate individual objects.

In image resolution, what's the difference between 400 dpi true and 400 dpi interpolated?

True resolution is the number of dots per inch of reflected light captured by the scanner head. The scanner head is physically limited in the number of dots which it can capture.

Interpolation software (usually built in to the scanner, but also available in raster manipulation programs) can make a good guess at what lies between the dots recorded by the scanner.

By comparing adjacent dots captured in a true 200 dpi scan interpolation software might replace each dot with four new ones (creating a 400 dpi interpolation) or even nine new ones (creating a 600 dpi interpolation).

For many applications (including the majority of paper to CAD conversions) interpolation works well. But interpolation will never be as accurate as a true resolution.

Is there an international standard for vectorization?

No. But conversion of paper to CAD and indexed raster images may provide you and your organization with an excellent opportunity to begin implementing ISO 9000 quality standards.

A good engineering document management system goes a long way towards meeting the ISO standard. See Trix Organizer.

I've too many drawings to vectorize in-house. How should I choose a conversion bureau such as the one run by Trix Systems?

There is a pretty clear relationship between price and quality in bureau conversion. We recommend starting with a pilot project comprising two stages. In the first stage establish that the bureau can deliver to your required quality for a small sample of drawings, say five or ten.

From this limited experiment set the parameters for a second stage test with a much larger sample of your drawings (maybe 5% of your stock). The second stage determines how your bureau will perform under a real load.

You and your bureau can establish a solid relationship during this test and sort out potential problems. And most important, involve your eventual users in the test. Have them check that the quality meets their needs. Some users may need help in defining their needs. You might want to think about an alternative to vectorizing all of your drawings. That is to scan all of your drawings and vectorize later as required. Trix Systems does offer quality-guaranteed Trix Systems Conversion Services.

.

If I have a small or hand-held scanner, can I scan a big drawing in pieces and 'stitch' them together afterwards?

If you have access to a large format scanner use it instead of stitching small images together. TracTrix can convert very large drawings (E/A0 and above). Many areas have copy shops or similar reproduction shops where a large document can be scanned to disk without editing at very reasonable prices.

For instance a busy 'E' size drawing scanned at 300 dpi monochrome (black and white) can be stored on a 3 1/2" diskette or emailed to you from your scanning bureau. The raster file on disk can then be vectorized in TracTrix at your convenience. Do be sure that your bureaus understands the level of quality and resolution you require before they scan the drawings. Vector output is only as good as the raster input.

Advice on how to scan for best results is included in TracTrix manuals.

What should I be aware of before I start converting images to CAD?

Vectorizing packages do not add information to your original drawing. We tend to think of drawings as being super accurate. This is not true. An image on paper is imperfect, inaccurate and imprecise. This may seem like a heresy to all who (like us) have slaved over drawing boards. Think about what we do when we 'read' a drawing:

Our intellect works wonders piecing together the disparate pieces of information which our eyes pick up from a drawing; Our brain automatically sorts text from line work and combines different sorts of visible information into a comprehensive whole. With this we can mill our piece, set our forms, etc. to the designer's requirement.

Contrast this with CAD. When we work in a CAD package vector information is created. The dimensions are the primary information and the CAD package produces a visible representation on screen (or to a plotter) based on these dimensions. This output is an approximation from the absolute dimensions of CAD. We go through the same processes to begin milling or setting up our formwork. And whether the information began in CAD or on a drawing board, we adjust it to the accuracy needs and limitations of the real physical world.

What are my options when considering how to get from paper to CAD?

There are several methods that will help you take a paper drawing or print into CAD. You could simply redraw from scratch in a CAD program, you could create a vector file from the drawing on a digitizing tablet, you could scan the drawing, open it in a 'Head-up' digitizing program on your monitor and digitize to vectors with a mouse much like the digitizing tablet

or you can use an automatic raster-to-vector program like TracTrix.

Tablet and head-up digitizing is subject to the skill and eye-hand coordination of the operator. Studies have shown that as digitizing operators tend to wander off the original lines by as much as 1/32 of an inch as they progress through a days work. Automatic raster-to-vector conversion will provide you with editable vectors in a fraction of the time of other methods.

So what's the process involved in translating paper to CAD?

It always requires two stages. In the first you scan and convert a raster image of the original drawing to CAD vectors . In the second, you check that the CAD representation created by vectorizing meets your need for accuracy and purpose.

You've probably heard the expression "garbage in, garbage out". This is not more true than scanning a paper drawing and vectorizing it. Vectorizing software can only make vectors where there is raster information for it to track.

Breaks in the lines, poorly connected lines at corners, blobs in the image at multiple line intersections and similar imperfections in the raster image all contribute to imperfect vectors.

TracTrix is used by a major prosthetics manufacturer to create cutting and drilling templates for microsurgery. Their originals are very finely detailed. Generally, the accuracy you'll achieve will depend primarily on the quality of the original drawing. Accuracy can then be made absolute in whatever CAD software you are using.

You should always expect to take raster images automatically to vectors and then edit them to the precision your application requires.

What else should I be aware of when considering paper to CAD conversion?

Manipulating large image files is best done with plenty of RAM.

For example, a 300 dpi scan of a B size (11 X 17) drawing will create an uncompressed file that is nearly 17 Megabytes. RAM therefore is an important consideration when converting and manipulating image files. At least 32 or even better 64 meg of RAM is recommended for paper to CAD conversion.

How about drawings that have been reduced to a smaller size?

When large drawings are reduced to a smaller size (using the 'reduce' function on a scanner or copier) lines and details tend to bleed together into indiscernible blobs of pixels. Parallel lines that are close together blend into one line. Small holes in the original drawing close up to a solid black object. TracTrix cannot 'read' through these solid black objects to figure out what they are. Therefore we do not recommend reduced size drawings as a starting point for automatic raster-to-vector conversion.

What is the difference between accuracy and precision in vectorizing

Repeating the point above: Scanning paper to create a raster image does not improve the quality in any way. As a matter of fact the scan is a less perfect than the original. Vectorizing software of any type can only vectorize that which it can 'see', the pixels in the raster image. The resultant vectors are no more perfect than the raster.

Vectorizing is accurate to the extent that the vectors are an accurate representation of the raster image. Is this 'accuracy' good enough for your application? Perhaps. But if it isn't, the vectors can be edited to make the vector file as precise as you require. Remember, vectors can be made to be mathematically perfect, rasters cannot.

Does high resolution (dpi) assure vector accuracy?

TracTrix follows either the centerline on lines or the outline of

images. It achieves this by looking for pixels. In centerline tracking, TracTrix looks for the chosen line width by looking at pixel quantities across line and seeks the mean center of the line. Having more pixels from a higher resolution will not necessarily create a "better" mean centerline.

When following an outline it looks for pixels on the edge of the image. Higher resolution might present a 'smoother' less jagged edge for TracTrix to vectorize but the vectors are not more accurately created.

Higher resolution makes for bigger files to manipulate. Bigger files require more pre-processing storage and processing takes much longer (twice the resolution means four times as many pixels to store and process).

As a general rule, there are rapidly diminishing returns from increasing resolution for raster to vector conversion. We tend to use 300 dpi or 400 dpi. Occasionally we need 600 dpi for small, finely detailed, originals.

Horizontal and vertical vector lines on screen do not appear to be straight, but they were in the original drawing. Why?

What you see on screen is only a visual display of the actual stored CAD file and does not represent what the file will create on a plotter or printer. The vector points, also seen on screen, more accurately depict the vectors as they will be plotted. What you are seeing is the result of a slightly askew scan.

When a computer depicts straight lines on screen they will appear straight, without a jagged appearance, only when they are absolutely square to the screen's matrix. That is if they are perfectly parallel to or at right angles to the horizontal scan of the monitor

If they are slightly off square the computer tells the screen to depict the straight line as a Cathode Ray Tube scan series which might be: Fill 100 pixels horizontally, go down 1 pixel, fill 100 pixels horizontally, etc.

It is this small one or two pixel step down which makes straight lines appear to have a jog or appear jagged on screen. The underlying vector which it is representing is a single point-to-point line.

Consider the jogs to be an optical aberration on the monitor screen and not representative of the actual vector. In TracTrix, one can see the actual vector points at each end of the line by clicking on the line.

How come some vector files are bigger than raster files of the same image?

The raster image size is based upon the resolution (dpi) and the physical size of the image. The vector file size is based upon the number of vector entities required to represent the raster file. The relationship between the two varies with the content of the drawings, the entity types in the vector file, the raster format used and whether there is compression in the raster format.

My drawings are already in CAD. How do I publish them electronically around my company?

After more than a decade of building engineering document management systems it is clear to us that raster file formats are the most secure, tamper-proof, durable and easiest method for publishing and archiving CAD designs. For some of the reasons take a look at our Trix RasterServer page.

In brief: Raster formats such as TIFF, CALS and C4 are non-proprietary, stable - you won't have to buy 'upgrades' every couple of years, and viewable and printable from a wide variety of viewing software.

How can I publish engineering drawings on the Internet?

The simplest method is to rasterize completed CAD designs from DWG or DXF to a raster format that can be viewed by a Web Browser. These formats are PNG, JPEG or GIF. These can

be published using the regular IMG tag in HTML. TracTrix and Trix RasterServer produce the PNG format for this purpose. The user needs nothing more than a recent Web Browsing software.

If you have large, complex, files and/or you want to do more than just view a drawing through a Browser you should look at adding a viewer plug-in to the users' Browsers. Plug-ins add functionality to a Browser. For example, Trix Systems offers a Plug-in version of Trix DrawingCenter. When this is installed a user can view DXF. DWG, HPGL and many raster file formats. The publisher does not have to convert these to PNG, JPEG or GIF. In addition to viewing the plug-in provides annotation, measuring tools, printing and saving controls. The user can add additional information to a copy of drawing and save it at the local workstation.

Other companies also offer plug-in viewers. There is an AutoCAD specific DWF format which also requires a plug-in. The PDF format is also for Web publishing using Acrobat Reader as the plug-in for Browsers. TracTrix and Trix RasterServer also convert CAD formats to PDF files.

Questions specific to TracTrix Software

What is the typical TracTrix Paper to CAD conversion process?

Scan your original into TracTrix. Select the description which fits your type of drawing from pre-defined types (e.g. architectural, contour, illustration, logo, mechanical). Select the portion of the drawing to convert. TracTrix then automatically processes the raster lines into vectors. Compare the vectors with the underlying original raster image and make any post-processing changes using TracTrix' built-in vector editing tools. Export the result in the vector format of your choice.

Speed/Expectations

TracTrix can convert a busy E-sized drawing in about 1 minutes on a reasonably modern computer.

In our years of providing image conversion for customers we have learned that individual needs are unique. So we are reluctant to suggest exactly what savings TracTrix can bring you. But based on the tens of thousands of drawings that we've converted in our own conversion bureau, we'd hazard a guess that after building in post-processing clean-up you could aim for 50% savings in hours spent on conversion.

That's not to say that all drawings lend themselves to raster-to-vector conversion. Depending on quality and intended application, you'll find some drawings can be almost entirely vectorized using TracTrix, others may inevitably require manual redrawing.

TracTrix is very forgiving in that the vectors created can be deleted and re-vectorized, edited and all vector errors corrected before saving in the CAD file format of your choice.

Will TracTrix give me a DWG file?

Yes.

TracTrix software also generates DXF, IGES and other vector description files which import into virtually all CAD packages. After you have imported the DXF, or IGES file into your CAD package you have a CAD document on your screen. When you open and save this from within whatever CAD program you have it creates the native CAD file for your CAD package.

.

Do I have to use TracTrix to drive my scanner, or will any image file work?

TracTrix can start with a raster image file from many different sources. Your, or your vendor's or your bureau's scanner(s) will produce a raster image file which TracTrix can work with. Digital cameras, scans from aperture cards and film are all raster files that TracTrix can convert. TracTrix also includes a TWAIN driver and large format scanner drivers to directly scan images into TracTrix.

What raster and vector formats does TracTrix work with?

Depending on version of TracTrix used, TracTrix can import ATT, BMP, BRK, CALS-Group 4 Type 1, CALS-Group 4 Type 2, CLP, CUT, DCX, DIB, GEM IMG, GX2, GTX G3, GTX G4, ICA, ICO, IFF, JEDMICS C4, JPEG, KFX, LV, PCX, MAC, MSP, PCD-Photo CD, PICT, PCX, PSD-Photoshop, RAS, TGA, TIFF- Pack Bits, TIFF-Modified Huffman, TIFF-Group III, TIFF-Group IV, TIFF-JPEG, WMF, WPG, XPM, XBM, XWD.

Depending on version of TracTrix used, TracTrix exports DWG/DXF (R12, R13, R14, 2000), DRW, EPS (AI-88), HPGL, IGES (4.0 and 5.1 - 112 or 126 entity), WMF and TIX formats.

How do I convert CAD files to non-CAD files for archiving or distribution?

TracTrix can individually convert DWG, DXF and HPGL CAD files to high-resolution raster files. For batch conversion consider Trix RasterServer.

Is TracTrix a scanner?

TracTrix is software. It does not include hardware. TracTrix can either drive or use the raster output from practically every scanner available. If you do not have a scanner you will need to acquire one to run TracTrix (unless you plan to use a bureau for your scanning). Plan to purchase a scanner which generates at least 300 dots per inch resolution and if a small format scanner preferably one that supports the TWAIN interface. TracTrix also drives several large format scanners including, Contex, Ideal, Vemco, CalComp, Océ and WideCom. You should also look for a scanner which also scans and outputs in color so that you can take advantage of TracTrix vectorization from colored originals.

I have an old scanner. Will it work with TracTrix?

If your scanner can create a monochrome (also called bi-tonal or black and white - i.e. no shades of gray) TIFF, CALS or one of many other monochrome image file formats you can use

TracTrix.

To use TracTrix color vectorization you'll need a color scanner.

Is there any way I can use TracTrix without a scanner?

The short answer is that TracTrix requires the type of image that scanners create. A more complete answer is that you do not need to own or acquire a scanner in order to use TracTrix. You could go to a scanning bureau to have your scans done. Or you could try using your fax modem, although we don't recommend this approach.

Conversion Bureaus: You could have your drawings scanned at a bureau, put the images on disk, tape or CD ROM and use these stored raster images as input to TracTrix.

A disadvantage of this approach is that you will not be able to experiment easily with scanning difficult images - dirty, dilapidated or cluttered ones - so that they vectorize optimally.

You and your bureau should establish clear quality control procedures for the scans, preferably after running a pilot study using a good sample of your drawings. Alternatively, for large conversion projects, you may want to consider contracting out all the work to a conversion bureau. Trix Systems does offer a conversion service. See Trix Systems Conversion Services.

Legacy Scans: If you already have stored binary images from previous scans, they can probably be taken into TracTrix with no problem. Some experimenting may have to be done to achieve this. For additional help with poor quality images ask us for advice.

Aperture cards: Few companies have an aperture card scanner. But many have aperture cards and an aperture card viewer. Aperture cards can be scanned by a bureau and the raster file delivered to you on disk or tape. These files can then be opened and converted by TracTrix.

Using a Fax image: We know of users who fax drawings to

their computer's fax modem (using the fine setting) and open the stored fax image in TracTrix to create vectors. However we don't recommend this approach. TracTrix works best with images of at least 300 dots per inch or 400 dots per inch. Normal fax standards do not support this level of resolution.

What will I save by using TracTrix?

Exact savings depend upon the quality of the original drawing and the type of image being brought into CAD. Some users report costs savings as high as 80% against conventional methods such as complete redraws, tablet digitizing and heads-up digitizing. And for some images, such as contours, automatic vectorization by TracTrix saves even more.

Taken over a library of original drawings we think you might realistically expect that the total time required to scan, vectorize with TracTrix and perform post-conversion tidy-up would be around 50% of what it would take for manual redrawing in CAD. However, every user's originals are different.

What will I gain by using TracTrix?

This question is a bit different from 'What will I save..? TracTrix has opened new methods to the 'time to market' issues so important to today's manufacturers. Many companies employ TracTrix to shorten time to manufacture by allowing sketches and other irregular drawings to be rapidly taken into CAD format for use in numerically controlled milling machines, machining centers, laser, water jet and plasma arc cutters.

Manufacturers are able to start with less than perfect originals and use TracTrix technology to produce samples and finished parts in a fraction of the time previously taken.

What does TracTrix do with text, symbols - Optical Character Recognition (OCR) ?

From our experience, text on engineering drawings, especially numerals and angled text, is very difficult to capture

adequately with conventional OCR technology.

Hand drawn letters and numerals often look alike and are very difficult for the eye to discern. For example, we find that the hand drawn 'B' and '8' are quite frequently mistranslated, a poorly formed capital letter 'G' is often mistranslated as a '6' and vice versa.

Text at any significant angle to the normal axes is missed entirely. Post OCR editing time to find and correct the things that were missed or worse, erroneously translated, often takes longer than keying in all the text. The comments for text apply equally to symbols.

TracTrix has a unique text recognition technology. When initiated it searches for text regions and isolates them from the drawing line work itself.

The text recognition in TracTrix will either convert the identified text regions to CAD characters by its built in OCR or allow you to 'train' it to recognize the uniqueness of the characters in your particular drawing or set of drawings.

If font-based text is not a requirement, text areas can be selected for vectorization using TracTrix outline mode. This produces very readable text in vector lines rather than a vector font.

Alternatively, if the text is too unreadable to convert automatically TracTrix provides tools to create font-based vector text. In TracTrix the user manually enters the text in vector overlay mode, then deletes the raster text underneath.

Trix Systems is continuously reviewing OCR and pattern recognition technology in order to improve and enhance TracTrix. Recognition of shapes is not yet available - this 'pattern recognition problem is similar to that encountered with OCR - but the techniques available for our present processors are just not satisfactory.

Can I just vectorize what I want and not the whole drawing?

Yes. TracTrix allows "windowing" around that portion you wish to vectorize.

Does TracTrix despeckle or eliminate "noise"?

Despeckling or noise elimination means eliminating small groups of pixels which create visual 'dirt' or visual clutter in a raster image. TracTrix settings allow the user to define the minimum size of object to be vectorized. In this way 'dirt' is ignored and not vectorized.

Does TracTrix do orthogonal adjustment?

Squaring up corners is referred to as orthogonal adjustment. TracTrix provides this facility. It also allows you to adjust the whole of or parts of a vector image to be absolutely vertical or horizontal.

Does TracTrix recognize circles?

TracTrix recognizes circles as well as arcs, lines, polylines and Bezier curves.

How far can TracTrix zoom in?

It can zoom enough so that the user can see down to the single pixel level.

Can TracTrix show a vector over a raster?

Yes, this is a facility in TracTrix and is a useful way to check that the vectors represent the raster image.

Can TracTrix show dimensions in metrics and/or inches?

Yes, either one, it's users choice each time it's used.

Can TracTrix be set to jump a gap in a line?

Incomplete or broken lines can be edited, in the raster with our raster editing tools or as vectors, into complete or single lines within TracTrix or as vectors in your CAD software.

How many layers can TracTrix put vectors on?

TracTrix can create and place vectors on a near unlimited

number of layers. This is achieved by identifying in advance the different raster line widths to be vectorized then each set of line widths is placed on different layers after vectorization.

TracTrix also recognizes offset spot colors. It will vectorize each different color and place the resultant vectors from each color on different layers. This too will provide an unlimited number of vector layers.

Can I use TracTrix without a CAD package?

TracTrix installs on your CPU as a standalone conversion tool that does not require a CAD package and as an object ARX application that runs within AutoCAD. No separate license is required. TracTrix 2 is a standalone package. Most users do have a CAD package into which the resultant TracTrix vectors are taken either before or after editing in TracTrix. Some NC users take their cutting input files directly from TracTrix.

How does TracTrix raster-to-vector conversion work?

An algorithm in TracTrix is designed to seek the pixels in the raster image and 'follow' them according to the parameters set by the operator. These include following the edge line of pixels in the 'Outline' following mode or following the centerline of the selected line width of pixels in the 'Centerline' following mode.

Additionally, TracTrix can be set optionally to create vectors as lines, arcs, circles, polylines, Beziers, NURBS or splines. The original image and desired file type often dictate the mode selected.

Definitions

DXF

DXF is one of many vector file formats used for the storage and exchange of CAD drawing data between different CAD software applications.

Hybrid file

A file containing both raster and vector information.

ISO 9000
ISO 9000 is an international certification of company quality standards established by The International Standards Organization (ISO), headquartered in Geneva Switzerland. Certification may be required in order to do business in certain industries, in certain countries, with certain companies and many government agencies. There are also other ISO standards with different numbers governing different types of businesses and technologies.

Line following
Line following software vectorizes interactively. The operator watches as the software tracks along the lines in an image and intervenes if an error is made. This is slower than automatic vectorization but can produce better results in certain types of complex images. Line Following will be supported in a future release of TracTrix in conjunction with other major technical advances.

Splines, B-splines, NURBS and Bezier curves
A spline or B-spline curve is a blended piece polynomial curve passing near a given set of control points. The blending functions are the type that provide more local control compared to other curves like a Bezier curve. A Bezier is a polynomial curve passing near, but not necessarily through, a set of given points.

A Bezier curve represents an equation of an order one less than the number of points being considered. A NURBS (Non uniform rational B-spline) is a B-spline curve or surface defined by a series of weighted control points and one or more knot vectors.

TWAIN
TWAIN is driver software which enables applications software such as TracTrix and other graphics applications to communicate with and drive your scanner. Once you've set up an application to use TWAIN, its presence should be almost invisible to you. Created by Hewlett-Packard, the TWAIN standard is now widely used by other small format scanner manufacturers.

If we haven't answered your question ask us here or call us on (978) 256-4445.

EXHIBIT

D

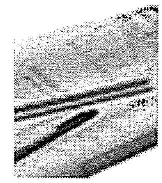
H-K Resources Pte Ltd

A Signature Wo

Drafting Services : **Architectural** **Civil** **Structural** **Mechanical** **E**

Introduction ¦ About Us ¦ Services ¦ Products ¦ Technology ¦ Contact Us

Why Convert Drawings to CAD?



There are traditionally three main ways of working on existing drawings, namely:

1. **Manually Maintaining Drawings on Paper**
2. **Scanning and Editing Raster Images**
3. **CAD Conversion**

CAD Conversion offers significant advantages over other methods including:

- ► **Reduced Cost of Revisions**
- ► **Increased Drawing Value**
- ► **Reduced Drawing Life Cycle Cost**
- ► **Reduced Storage And Creation of a Standard Filing System**
- ► **Ability to Enhance Company's Competitive Advantage**
- ► **Ability to Maintain Consistent Level of Quality for Projects**

1. Reduced Cost of Revisions

Revisions done in CAD may be 2-8 times faster than the same revision done by manual methods. And converting drawings to fully vectorised files (instead of just raster images) increases the ease of editing since the user no longer needs to switch back and forth between raster and vector files. Conversion allows the full capabilities of a CAD system to be utilised. And what's best, CAD conversions are low-cost, usually less than $50 per drawing.

Back to Top

2. Increased Drawing Value

Once drawings are in CAD format, the uses for it increase dramatically. For instance, intelligent CAD files can be used with cost estimating software, facilities management applications such as area calculation and inventory tracking as well as engineering design and analysis software or numerically controlled machining in manufacturing. Intelligent CAD drawings can also significantly reduce the time required to extract data from drawings and enter it into databases which are used for such things as project management, quality assurance, maintenance and material control. Time savings are estimated to range from a few days to even a few months for large projects.

Back to Top

3. Reduced Drawing Life Cycle Cost
Since conversion is a one-time cost, the second and third revisions to a drawing produce even greater savings, reducing the overall cost of maintaining a drawing throughout its useful life. This saving also results in earlier payback for conversion projects. In addition, since physical drawings tend to be huge, cost-effective storage become major issues if drawings need to be kept for a substantial length of time. Electronic CAD drawings can easily be kept on a single hard-disk.

Most building control authorities require professionals to maintain copies of all project drafts and drawings for a substantial period of time (sometimes as long as 10 years!!). For a medium-sized firm, compulsory storage of these physical drafts have been found to take up as much as 2,000 sq ft of space annually. That's enough space to house an additional 20 staff, or the equivalent of paying an additional $10,000 every month. All these can be stored electronically on a single box of CD-ROMs.

Back to Top

4. Reduced Storage Requirements and Creation of a Standard Filing System
Establishing CAD as the standard filing procedure will decrease the amount of engineering time spent looking for drawings, and also the number of lost drawings. Filing more drawings electronically may also reduce square footage being used to store paper drawings. Typically 9 filing cabinets (20,000 pages) or over 3,000 large format drawings can be stored on a single CD-ROM disc and any individual file, word, phrase or drawings can be located within seconds via anyone with appropriate network access.

Back to Top

5. Obtain a Competitive Advantage
Converting drawings to CAD allows a firm to project a consistent, progressive and high quality image to their clients by eliminating the use of outdated manual drafting methods. CAD is recognized as the industry standard in technical drafting and can now be used for drawings which were created before CAD.

Back to Top

6. Maintain a Consistent Level of Quality for all Your Projects
Begin the commitment of quality at the very first stage of any project, even those projects utilizing old and tattered paper drawings. With CAD conversion, company standards such as text fonts, line weights and other drafting standards are enforced. CAD drawings offer a much higher and more consistent level of quality.

What's more, it only costs as little as $50 to convert your paper drawings to CAD. So why not contact us now to find out more!

..... your one-stop drafting soluti

EXHIBIT

E

Examining the Encryption Threat

Jason Siegfried
Christine Siedsma
Bobbie-Jo Countryman
Chester D. Hosmer
Computer Forensic Research and Development Center

Abstract

This paper is the result of an intensive six-month investigation into encryption technologies conducted at the Computer Forensic Research & Development Center (CFRDC) at Utica College. A significant number of encryption applications were collected and cataloged. A roadmap for the identification of the unique characteristics of encrypted file formats was created. A number of avenues were explored and the results documented. The actual process is not outlined comprehensively due to proprietary needs; however, the following briefly details the process and the significance of our findings.

Introduction

In 2001, a firestorm of controversy erupted in the case of United States V. Nicodemo Scarfo Jr. At issue was the use of Carnivore, a covert key-logging tool that had been the subject of much scrutiny, and its sophisticated successor, Magic Lantern. Because the suspect used advanced encryption technology, law enforcement had to use a sniffing keystroke logging tool. The legal and covert deployment of carnivore and magic Lantern caused many law-abiding citizens to feel that the time of the Orwellian coined term, “Big Brother” had arrived. However, it became evident that law enforcement was unable to decrypt and access encrypted data. The Scarfo case concerning law enforcement’s need for such tactics as Carnivore or Magic Lantern produced fear in law abiding citizens and demonstrated that law enforcement did not have, nor currently has, a better option.

Law enforcement is currently at the mercy of criminal or terrorist entities that employ sophisticated encryption applications. The future success of Magic Lantern is questionable considering two factors: 1) law enforcement must be aware of criminal activities prior to installing the Magic Lantern tool; and 2) the hacker community will not allow such covert techniques to persist, as evidenced by the following quote obtained via Google’s cached feature from a website that is no longer available on the Internet,

> Seeing as how some antivirus software manufacturers will not be looking for the FBI’s Magic Lantern virus, it seems to me that the open source/free software community should be doing what it does best: doing it ourselves.[1]

[1] Investigating Cyber Knight. Posted 24 Nov 2001 by Pseudonym. Original URL <http://www.advogato.org/article/384.html> is no longer available, but access to

The hacking community's ability to defeat new technologies jeopardizes the success of Magic Lantern.

The progressive sophistication and strength of encryption technologies remains a significant obstacle to law enforcement efforts to obtain digital evidence protected by sophisticated mathematical manipulations. The strength of encryption applications consistently advances; the number of encryption applications continues to multiply, and the availability of these sophisticated applications via the Internet continues to increase. Regardless of the grandiose speeds of modern computing technologies, the ability to crack sophisticated encryption tools employed by criminal or terrorist entities remains mind-boggling. The following table demonstrates the machine power required to crack an encryption key in 1997.

Encryption Name & Strength	Time Taken to Crack Key	Machine Power Required to Crack Key	Maximum Speed Required to Crack Key
48 bit RC5	13 days	5000 max, 7000 overall	440,000,000 keys/sec
56 bit RC5	270 days	4000 teams, 10,000's machines	7,000,000,000 keys/sec
64 bit RC5	1,470 days	Not Available	88,000,000,000 keys/sec
Elliptic Curves (109 bit)	120+ days	9,500 in total, 5,000 active at one time	Not Available
RSA 512 bit	Polynomial selection – 2.2 months Factoring – 5.2 months	292 plus a Cray for the last stage	Not Available
56 bit DES	~90 days	Max: 14,000 in a single day	7,000,000,000 keys/sec

Table 1 – Required Time, Machine Power, and Speed in 1997 to Crack Encryption[2]

While 1997 data may seem outdated, the correlation of increasing encryption keys consistently increases along with computing power. In 1997, did law enforcement have the type of machine power, manpower, or financial support to devote such resources to cracking one single encryption key? How likely is it that law enforcement has the resources today to crack the encryption keys deployed in 2004? Furthermore, as the term "quantum encryption" is appearing in security conferences and underground hacker sites alike, law enforcement's ability to catch up to sophisticated encryption tools is nil.

Encryption applications have historically been deployed for legitimate purposes such as privacy, protection, and security. However, the utilization of advanced encryption

<http://216.239.37.104/search?q=cache:6EXloJTwLakJ:www.advogato.org/article/384.html+Investigating+Cyber+Knight&hl=en&ie=UTF-8> is available.

[2] Brute force attacks on cryptographic keys. <http://www.cl.cam.ac.uk/~rnc1/brute.html>. Accessed 21 January 2004.

algorithms has developed into a dual technology applied for legitimate as well as nefarious purposes. In 1997, Dorothy Denning and William Baugh made the following statement, "...our findings suggest that the total number of criminal cases involving encryption worldwide is at least 500, with an annual growth rate of 50 to 100 percent." [3] With the ease of use, current availability, and multiple hacking communities, it can be presumed that even Denning and Baugh understated the use of encryption technologies by criminal and terrorist entities. In the 1999-2000 document, Current U.S. Encryption Regulations: A Federal Law Enforcement Perspective, the author describes the threat as follows.

> ...Absent some form of key recovery or recoverable method, a brute force attack will not meet law enforcement needs. If we are working on a terrorist case and intercept a communication that we believe to be in furtherance of criminal activity, and that communication is encrypted – say with PGP, which is 128 bit encryption, a brute force attack to decode one PGP message, using a Cray computer, would take nine trillion times the age of the universe... This is our greatest fear, that, one day, a terrorist attack will succeed because law enforcement could not gain immediate access to the plaintext of an encrypted message...[4]

Without the use of a covert key logging technology such as Carnivore or Magic Lantern, the use of sophisticated encryption applications can stop a digital investigation cold in its tracks. Encrypted data has become a clear obstacle to the furtherance of successful computer forensics investigations. This paper details an intensive six-month research effort, which identified a number of significant characteristics that can be incorporated into a digital forensics investigation. It is hoped that it will provide a number of benefits to law enforcement professionals.

The ability to identify encryption applications using forensic file identification techniques is one that has not yet been seriously explored. Although this six month manually intensive study did not produce an easy way to expedite the cracking of an encryption key or password, it certainly did produce a number of significant results that will expedite the identification of the utilization of an encryption application, among other characteristics of the encryption application.

Currently, random, unintelligible data, not immediately attributed to a file can be inadvertently identified as binary file remnants, previously deleted data, or partially overwritten files, while in fact, it is possible that remnant data can be attributed to encrypted data. The significance of this study's findings can support and assist investigators in quickly identifying the presence of an encryption application, the specific

---

[3] Dorothy Denning and William Baugh. "Encryption and Evolving Technologies as Tools of Organized Crime and Terrorism."

[4] Smith, Charles Barry. 1999-2000. Current U.S. Encryption Regulations: A Federal Law Enforcement Perspective. <http://www.law.nyu.edu/journals/legislation/articles/vol3num1/smith.pdf>. Accessed 21 January 2004.

encryption application used to encrypt digital data, and the signature and/or patterns associated between the encryption application and its subsequent encrypted data.

File Identification through Binary Analysis

A file header is the first portion of an electronic file that contains metadata, as opposed to data.[5] "Metadata is the background information that describes the content, quality, condition, and other appropriate characteristics of the data."[5] It is essentially "data about data." The file header itself is transparent to the user and can only be viewed with a low-level disk viewer/editor. It contains information necessary for the application to "recognize" and "understand" the file. The presence, byte size, and data content of file headers are unique to virtually every application. For example, a Microsoft Word document (.doc) contains very structured and lengthy headers and footers embedded throughout the file (10,752 bytes), as opposed to a basic text file (.txt) that does not even have a header or any other embedded data. Although file header content varies from application to application, the most consistent feature is the presence of a file signature.

File signatures, unlike file extensions, are not easily altered and thus the more accurate means of file identification. Additionally, file extensions are generally limited to only three or four characters; the extension itself tends to be reused for multiple file types.[6] Forensic file type identification is a process used by computer forensic investigators to examine the metadata that applications embed in the files that they create (file header and/or footer), and is the most reliable way of identifying the actual file type. Like any other application that creates files, it is assumed that the resulting encrypted file will have embedded metadata that the file encryption application would use to recognize it as "one of its own," not just by the file extension, but also, the addition of file header and/or footer information.[7]

One purpose of this study was to advance forensic file type identification to the next level through very deep and low-level analysis of encrypted files. The goal for this phase of the experiment was to expand the scope of research to identify not only file signatures, but other important metadata as well. The result was a process to recognize encrypted file signatures and extract detailed information from the encrypted file header.

Two popular file encryption applications were chosen to perform the deep, low-level analysis on. Two programs were chosen to achieve some diversity: RipCoder,[8] very

---

[5] http://inside.uidaho.edu/tutorial/overview/overview.htm

[6] As an example, the .doc extension; commonly recognized as the extension for Microsoft Word documents, a file with that extension could possible one of nine other known file types. See http://www.filext.com/detaillist.php?extdetail=doc

[7] Commonly referred to as 'file signatures.' For a sampling of file types and their associated file signatures, see http://www.garykessler.net/library/file_sigs.html

[8] RipCoder's homepage, http://kach.nm.ru/

basic, easy to use program and FineCrypt,[9] an advanced one with many user-defined options. These popular software programs were obtained freely and anonymously from the Internet. As can be seen from the illustration below, the webattack.com download site had FineCrypt listed as the featured download with RipCoder appearing as well.[10]

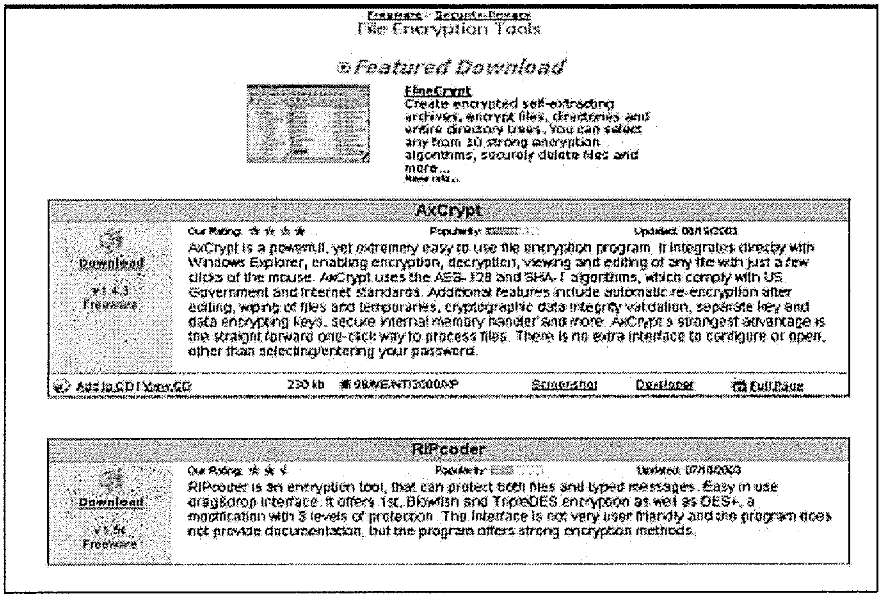


Figure 1 – Screenshot from webattack.com Download Site

Experiments were conducted by encrypting files from a standard dataset with combinations of user-defined parameters that are unique to virtually every application. The test dataset consisted of one, two, and eight-byte text files (.txt) along with a 256-byte binary file with each byte representing a different ASCII character starting with the hexadecimal value 00, and ending with the hexadecimal value FF. As the number of options increase with more advanced software, so too does the number of permutations of settings that must be tested. (The FineCrypt analysis required the production of more than 640 encrypted files.)

[9] FineCrypt's homepage, http://www.finecrypt.net/
[10] webattack.com's homepage, http://webattack.com/

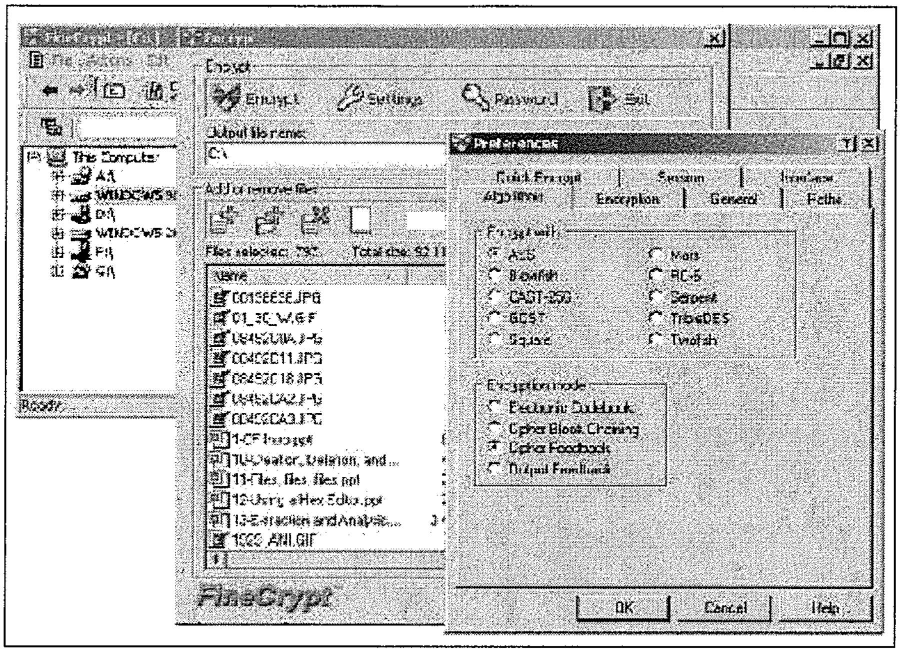


Figure 2 – FineCrypt Interfaces

The resulting encrypted files were then analyzed with a low-level disk viewer to identify metadata contained in the headers and footers of those files. The values in the headers of these files were examined as single byte and byte block values. The key to successful pattern analysis lay in the ability to identify the static header structure and associate the dynamic values with specific attributes of the unencrypted file and/or user-defined options. In addition to the test dataset, a number of files ranging from zero to several thousand bytes were created, encrypted, and analyzed at the experimenter's discretion to pursue predictable value patterns. In order to successfully and efficiently manage and track a dataset of that magnitude, a naming convention using fields based on user-defined options was established. The naming convention allowed for quicker comparisons between encrypted file characteristics and the resulting header values. The following illustrations are screenshots of RipCoder and FineCrypt files as seen with a low-level disk viewer.

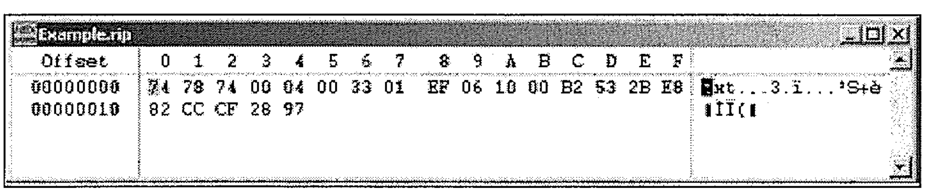


Figure 3 – RipCoder File in Low-Level Disk Viewer

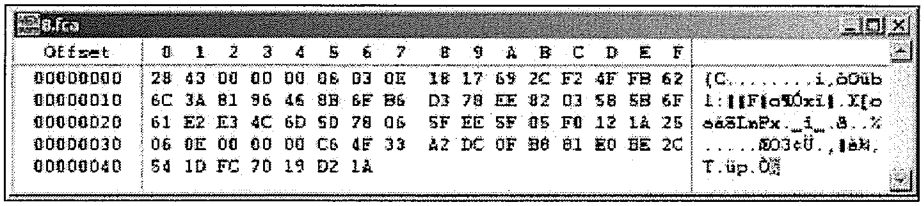


Figure 4 – FineCrypt File in Low-Level Disk Viewer

The analysis efforts were extremely successful. Significant details and characteristics of the unencrypted and encrypted payloads were identified through rigorous examination and analysis of the encrypted files and file headers. The following information can be *located* and *extracted* from the metadata contained in the above files:

- Application signature for positive program identification
- Encryption algorithm used to encrypt payload
- Encryption mode used to encrypt payload
- Password (yes/no) and location of password byte block data
- Key (yes/no) and location of key byte block data
- Compression (yes/no)
- File extension of unencrypted file
- Number of characters in unencrypted file name and location of the bytes representing the name (varies with size of name)
- Encrypted file size excluding four-byte checksum (location of checksum bytes was discovered)
- Number of bytes of cipher text and exact location
- 32-bit write-back option for DES+ algorithm (yes/no)

As an example, consider the FineCrypt header below and note the hexadecimal value of the highlighted offset.

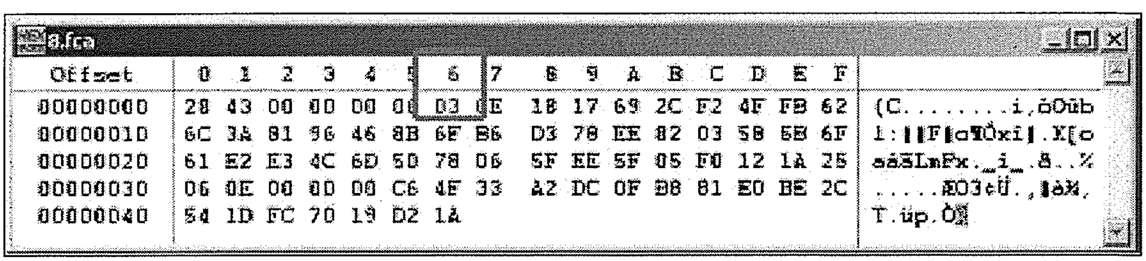


Figure 5 – FineCrypt File in Low-Level

The hexadecimal value of 03 indicates that the algorithm used to encrypt the file was AES and the encryption mode employed was Cipher Feedback. The value of offset 6 will always represent the algorithm and mode selection in FineCrypt files. The complete hexadecimal value matrix for offset 6 appears in the following table.

Offset 06					
Value	**Mode**	**Algorithm**	**Value**	**Mode**	**Algorithm**
00	?????????????????	??????????	15	Electronic Codebook	MARS
01	Electronic Codebook	AES	16	Cipher Block Chaining	MARS
02	Cipher Block Chaining	AES	17	Cipher Feedback	MARS
03	Cipher Feedback	AES	18	Output Feedback	MARS
04	Output Feedback	AES	19	Electronic Codebook	RC-6
05	Electronic Codebook	Blowfish	1A	Cipher Block Chaining	RC-6
06	Cipher Block Chaining	Blowfish	1B	Cipher Feedback	RC-6
07	Cipher Feedback	Blowfish	1C	Output Feedback	RC-6
08	Output Feedback	Blowfish	1D	Electronic Codebook	Serpent
09	Electronic Codebook	CAST-256	1E	Cipher Block Chaining	Serpent
0A	Cipher Block Chaining	CAST-256	1F	Cipher Feedback	Serpent
0B	Cipher Feedback	CAST-256	20	Output Feedback	Serpent
0C	Output Feedback	CAST-256	21	Electronic Codebook	3DES
0D	Electronic Codebook	GOST	22	Cipher Block Chaining	3DES
0E	Cipher Block Chaining	GOST	23	Cipher Feedback	3DES
0F	Cipher Feedback	GOST	24	Output Feedback	3DES
10	Output Feedback	GOST	25	Electronic Codebook	Twofish
11	Electronic Codebook	Square	26	Cipher Block Chaining	Twofish
12	Cipher Block Chaining	Square	27	Cipher Feedback	Twofish
13	Cipher Feedback	Square	28	Output Feedback	Twofish
14	Output Feedback	Square			

Table 2 – Offset 6 Signature Values

The file header structure and value associations remained consistent regardless of the unencrypted file type. Additional tests were run using Microsoft Word, Power Point, and Excel files. Image files were also considered and tested to ensure consistency (.jpeg, .gif, and .bmp). The structures and values remained consistent with very large binary files as well (600 MB random binary file.)

Additional Testing

The deep, low-level analysis of these two file encryption applications produced a significant amount of data. The additional phases of testing involved monitoring file and registry activity during encryption, examining slack space, swap space and unallocated space for passwords and encrypted file content, byte boundary analysis of encryption

algorithm and mode padding schemes, and finally, identifying and locating files and registry keys that remained on the test computer after uninstalling the application. A brief discussion of the install/uninstall monitoring results follows.

While RipCoder is a stand-alone executable and does not require installation because it runs from its own program folder, FineCrypt requires its system files to be installed on the computer. We developed a process using installation monitoring software and a text comparison utility to capture and analyze all file and registry activity during installation and uninstallation of applications. The table below summarizes the installation results.

FineCrypt Installation	Files	Registry Keys
Added	48	672
Modified	5	24
Deleted	8	32

Table 3 – FineCrypt Installation Data

After the application was uninstalled, 118 registry keys and eight (8) files remained on the computer. After the system was rebooted, all 118 registry keys remained, but only one of the eight (8) files was present. Although RipCoder runs as a stand-alone application, two ".rip" folders were created in the registry and remained even after the program was deleted from the system. After uninstalling and deleting these applications, file and registry remnants resided on the system as conclusive evidence of prior existence.

Conclusion

Enabling law enforcement to easily identify encrypted files on a suspect machine is only the beginning of what should be continuing research efforts. Although the probability of developing a unique process to easily crack encryption keys or passwords remains quite unlikely, the significant findings produced by these research efforts suggest that small steps can be taken to assist and support law enforcement efforts in analyzing and extracting critical digital evidence in the presence of an encryption application. This research effort produced several significant outcomes. The following are the accomplishments to date.

- Encryption applications were collected and cataloged, establishing a large data set on which to conduct further analysis (455 applications).
- Using this collection, a database of hash values was created (10,529 files), as a tool to aid in the forensic identification of encryption applications.

- Processes and procedures were developed for the identification and extraction of encrypted file metadata.
- Processes and procedures were developed for all other phases of testing including, but not limited to, application remnant identification, system monitoring during encryption, swap and slack space analysis, and cipher text padding analysis.
- A geographical study was launched into the origins of current encryption technologies.
- A roadmap was laid for continued research into the area.

It is imperative that research and development efforts continue to advance the innovative solutions available to law enforcement to combat the strength of modern and continuously progressive encryption applications. The findings produced by this research effort significantly mitigate the time consuming processes of manually identifying encryption applications and what encryption algorithms were used. As research continues, the potential to overcome the impressive leads that criminal and terrorist entities currently maintain with the use of encryption could be significant, without the need to work against the law-abiding public.

For information on obtaining a complete copy of the Encryption Report, please contact Christine Siedsma at the Computer Forensics Research and Development Center. (CFRDC) csiedsma@utica.edu

About the Authors

Chet Hosmer (chet@wetstonetech.com) is a co-founder, President and CEO of WetStone Technologies, Inc. He has over 25 years of experience in developing high technology software and hardware products, and during the last 11 years, Chet has focused exclusively on information security technologies. This focus has resulted in technology innovations in secure time stamping, steganography, network and host based intrusion detection systems, digital watermarking and digital forensics. Chet is a co-chair of the Technology Working Group, one of the seven working groups of National Cybercrime and Terrorism Partnership Initiative sponsored by the National Institute of Justice. He is the Director of the Computer Forensics Research and Development Center (CFRDC) of Utica College, and holds a B.S. degree in Computer Science from Syracuse University.

Jason Siegfried is a Computer Forensics Specialist at WetStone Technologies, Inc. He specializes in Computer Forensics and Information Assurance research. His most recent work includes investigating, collecting, and documenting digital forensic tools and cyber weapons. He completed his internship while working on the Encryption Threat project at the Computer Forensics Research and Development Center during the

summer of 2003. He joined WetStone Technologies, Inc. in the fall of 2003. Jason graduated from the Economic Crime Investigation program with a concentration in Computer Security from Utica College of Syracuse University. He graduated with summa cum laude honors in receiving his B.S. in Criminal Justice from Utica in August of 2003. Prior to completing his B.S. degree, he graduated from M.V.C.C. with an Associates degree in Liberal Arts and Science.

Bobbie-Jo Countryman is a Digital Forensics Analyst at WetStone Technologies, Inc. She specializes in the research and investigation of various Information Assurance and Digital Forensics topics. Her most recent project involves a research and investigation effort into the cyber-technologies deployed in America's Critical Infrastructures. She currently acts as a team leader to produce in-depth, detailed reports to expose potential vulnerabilities within America's Critical Infrastructures that provide opportunities to hostile adversaries. Additionally, she assisted research efforts to investigate a number of biometric devices, their capabilities for integration, and their potential for Information Assurance security. She recently earned two Associate of Science degrees consecutively, for Computer Information Systems and Computer Forensics.

Christine Siedsma is the project Manager of the Computer Forensic Research & Development Center at Utica College, where she has supported a number of research efforts in the fields of Computer Security and Digital Forensics. She is also an adjunct instructor of Computer Forensics, and maintains the E-Evidence.info website. Christine earned her B.S. degree in Criminal Justice from Utica College of Syracuse University.

Project Team

The following organizations and individuals contributed their time, effort, and expertise to complete this project:

Air Force Research Laboratory (AFRL)
Joseph Giordano, Peer Review Committee

Northrop Grumman TASC
Dan Kalil, Peer Review Committee

Syracuse University
Dr. Shiu-Kai Chin, Peer Review Committee

Computer Forensic Research and Development Center (CFRDC)
Chester D. Hosmer, Director
Christine Siedsma, Program Manager
Jason Siegfried, Intern
Sabina Bajic, Intern
John Sallustio, Intern

EXHIBIT

F

Interactive Numeric & Spatial Information Data Engine

Why Should I Create Metadata?

This sub-page of Metadata topic list some reasons why metadata should be created.

There are many reasons to create metadata. Metadata servers numerous important purposes such as data browsing, data transfer, and data documentation. Here are some additional benefits to think about:

- Metadata helps users answer questions about the data.
- Metadata helps publicize and support the data you or your organization have produced.
- Metadata supports the creation of a data inventory. Documenting data and its availability provides agencies with the means to measure production.
- Metadata that conform to the FGDC standard are the basic product of the National Geospatial Data Clearinghouse, a distributed online catalog of digital spatial data. This clearinghouse will allow people to understand diverse data products by describing them in a way that emphasizes aspects that are common among them.
- Metadata may be considered insurance. Having metadata available insures that potential data users can make an informed decision about the appropriate use of a data set.
- Metadata is a key component of data lineage. It provides basic information about the source and derivation of a data set.

The state of Idaho has realized the importance of geospatial metadata by the creation of the following standards and guidelines:

Category: S4220 – Geospatial Metadata

Category: G320 – Geographic Metadata Guideline

< Previous Tutorial Topic Next Tutorial Topic >

EXHIBIT

G

Pinpoint Labs Blog

How to Preserve, Collect, Recover and Filter Electronic Evidence

---

« Recovering deleted files (3 of 3)
Understanding File Timestamps »

What are File Headers? (Signatures)

Many file types can be identified by using what's known as a file header. A file header is a 'signature' placed at the beginning of a file, so the operating system and other software know what to do with the following contents.

Many electronic discovery applications will use the file header as a means to verify file types. The common fear is if a custodian changes a files extension or the file wasn't named using an applications default naming convention, that file will be missed during electronic discovery processing. For example, if I create a Microsoft Word document and name it 'myfile.001', instead of 'myfile.doc' and then attempt to locate all Microsoft Word files at a later date, I would miss the file if I were looking for all files ending in '.doc'. There are specific file extensions associated with the native application.

During a computer forensic investigation file headers are extremely valuable because they allow us to locate the contents of deleted files, user activity logs, registry entries, and other relevant artifacts. For example, if I'm investigating a custodian hard drive for evidence that they were working for a competing company I would want to recover their file activity records. A large number of custodian activity records are often already purged or deleted. By scanning a computers hard drive for the signature related to user activity records we often recover relevant artifacts (file access records) up to several years after they were deleted.

Print This Post Email This Post

Tags: artifacts, file headers, file signatures, unallocated space

This entry was posted on Tuesday, August 12th, 2008 at 2:22 pm and is filed under Data Recovery, File Header, Tips & Tricks. You can follow any responses to this entry through the RSS 2.0 feed. You can leave a response, or trackback from your own site.

Leave a Reply

You must be logged in to post a comment.

EXHIBIT

H

WHITE PAPER

WHY GROUP 4 TIFF/CALS
IS THE
FORMAT OF CHOICE
FOR ELECTRONICALLY DISTRIBUTING
AND ARCHIVING DRAWINGS
USED IN
IN THE CONSTRUCTION INDUSTRY

This paper examines two formats that are often put forward as alternatives for electronically distributing construction drawing documents in applications such as permit approval, bid solicitation and facilities management. The paper discusses Group 4 TIFF/CALS and PDF: their origins and structural characteristics; and presents the file storage requirements of a typical original CAD document converted to each format to provide a frame of reference for the discussion of the relative benefits of each format type.

Introduction

This White Paper discusses the Group 4 TIFF (Tagged Image File Format) or CALS (Continuous Acquisition and Life-cycle Support), and PDF (Portable Document Format) format types for use in distributing drawings in the construction industry for Electronic Bid Solicitations.

The main reason such a choice is necessary is that it is generally accepted that distributing original CAD design documents is impractical due to their proprietary nature, their lack of inherent 'original work' protection for the originator and their typically large file size.

Industry has dealt with this distribution issue in the past mainly by producing paper copies of the documents. However, as the inefficiencies of this method have become increasingly obvious, alternative formats have been put forward as the means of performing such distribution electronically by either converting the original CAD format or by scanning the paper copies and outputting to a particular format.

In examining this subject, it is easy for confusion to arise or misinformation to occur through over-simplification of what is a relatively complex issue. For example Group 4 TIFF is often confused with other variants of the TIFF format, which produce greater file sizes than Group 4. This paper attempts to clarify key issues sufficiently without delving into an overly technical discussion.

Origins and Structures

Group 4 TIFF is an 'open' industry-standard raster* format that was designed by CCITT as general monochromatic format for use in the copying and facsimile industry where compactness and image versatility are primary considerations. Group 4 TIFF is part of the TIFF format, which has several different variants using different compression techniques. Group 4 is widely used in the reprographic industry and employs a very efficient compression capability.

The CALS file format is a US government-accepted variant of the Group 4 TIFF specification specifically developed and supported as an archival standard within the government. CALS is almost identical to Group 4 TIFF except for header information. For purposes of this document these formats are therefore regarded as being virtually the same.

PDF is a proprietary file format controlled by Adobe Systems and in certain situations has license cost implications. It is commonly used to convert disparate word processing and graphic application files into a document format that can only be viewed with the PDF viewer called Acrobat Reader. A PDF document contains one or more pages consisting of text, graphics and images produced directly from applications or from files containing PostScript page descriptions. A PDF document may also contain information in electronic representation only, such as hypertext links.

Unlike Group 4 TIFF which stores information and perform compression in a uniform way PDF can be a composite structure containing multiple format types and compression techniques. To reduce file size, PDF supports multiple compression filters including: JPEG compression of color and grayscale images; CCITT Group 3, CCITT Group 4, LZW and Run Length compression of monochrome; and LZW and Flate compression of text, graphics and indexed image data.

In "automatic" compression mode the PDF creation tool analyses the document structure and then chooses what it deems appropriate.

File Sizes

The first thing to note is that, because of the way each of these file formats work, it is not possible to develop an exact formula or relationship between format types for the specific amount of storage space that will be used for a particular drawing. How much space each format type will take to represent a particular drawing is dependent on: the attributes and complexity of the drawing, the size of the drawing and the resolution level at which you wish to store the document. The ratios applicable to each format type therefore is not consistent from document to document.

To provide a frame of reference, however, the following is an actual example of a typical drawing document:

Original CAD format (DWG)	-	1.62 megabytes (MB)
PDF (using automatic default compression)	-	24.5 KB
Full Resolution (200 dpi) TIFF Group 4 (or CALS)	-	13.3 KB
Preview Resolution TIFF Group 4(or CALS)	-	3.3 KB

Compression Accuracy

Unlike PDF, which uses a variety of compression techniques (as described above) including lossy (which 'throws-away' selected data), the TIFF format uses strictly non-lossy compression algorithms. While PDF complex compression capabilities work well for complex documents, this adds overhead to relatively uniform document structures such as construction drawings. In addition, lossy compression can result in inaccurate reconstruction/scaling of the images which can render them unreliable for scaled printing and for calibration to perform electronic on-screen measurements: an important function in the take-off process needed with bid documents.

Internet Distribution

The ease of access to documents and the options available for their download and use off-line are major areas of consideration when they are being distributed over the Internet.

TIFF/CALS - A technique used to further improve the access speed of raster drawings over the Internet is to provide a lower resolution preview image in addition to the full resolution image for a particular document. Users can browse in preview mode, select the documents they wish to download and then download the full resolution documents in an unattended batch mode at a later time. The ability to selectively view and then download individual TIFF/CALS images from within a structured document is extremely important to users.

PDF – With PDF, multiple pages are stored as one document and cannot be retrieved online individually as with the TIFF example. The extended time needed to download the complete PDF file in order to work with any one part of the document off-line adds substantial overhead and makes the process extremely cumbersome especially with large projects.

Security

Design professionals have long been comfortable with paper distribution for which TIFF or CALS is the electronic equivalent and provides the greatest amount of 'original work' protection. TIFF/CALS cannot be easily imported into a vector format and therefore is well suited for public distribution.

Document Organization & Tools

The organization and presentation of TIFF drawings is performed by an application such as MaxView that was designed to handle TIFF/CALS documents. MaxView is specifically oriented towards construction document handling and its functionality, for both the author and the viewer of these types of documents, is far superior to that of the PDF equivalent (Acrobat).

Acrobat only allows the user to import drawings in a variety of sizes and formats up to and including E-size; and Acrobat makes no provision to support other format types within its organization structure. PDF documents have basic navigation capabilities but have no count and measurement capabilities for takeoff functions.

MaxView organizes all the document of a project into an intuitive tree-structure with folders and files that maintain their original document format. Typically, drawings are stored in TIFF/CALS while specifications are stored in TIFF, Word, Excel or PDF format. MaxView provides the user with a way to calibrate the TIFF

drawing to a specific scale and then to complete takeoffs using the count, distance and area measuring tools intrinsic to the MaxReader application. This functionality coupled with the ability to accurately print any drawing size to scale are major advantages for the use of MaxView in plan review and electronic bid solicitation applications.

Other Considerations

There are a number of other factors that impact the format type to be selected:

- In many situations, it is necessary to convert legacy documents from a paper source to an electronic format. TIFF images are the standard for scanned drawings.
- Since TIFF is so widely accepted and used in the reprographics industry, it is very easy for end users to obtain printed paper copies if they desire.
- Major US government organizations such as USACE and the Air Force have standardized on the CALS format; giving it assurance as a viable long-term archival vehicle.

To summarize; Group 4 TIFF/CALS has a size advantage over PDF and with the use of previews, this advantage is extended. Combining this with better 'original work' protection; its accurate scalability; its ability to work with all forms of document inputs including paper; its 'open' non-proprietary nature and general community acceptance along with a better selection of available tools gives TIFF/CALS a decisive edge as the format of choice for distributing electronic plan pages used in the construction industry.

About the Author: Key Churchill is a document imaging expert who, as a principal of Integrated Imaging Inc, has been involved with many imaging application developments including: design and implementation of the Valley Construction News website in Roanoke, Virginia which has performed electronic plan distribution since 1997 and more recently MaxView's MaxPlans website.

EXHIBIT

I

Electronic Records: Preservation and Access

October 6, 2005
Dr. Charles Dollar

Seventh in the *Missouri Electronic Records Education and Training Initiative (MERETI)* workshop series

- Download Workshop On-Screen Presentation (13.3MB)
- Download Workshop Printed Handouts (1.32MB)

Note: *Click on the to watch the instructor discuss key points. The number refers to the corresponding slide in the accompanying PowerPoint presentation and handout.*

In this advanced workshop, Dr. Dollar explores the many elements involved in the long term preservation and use of electronic records. He explains what digital archiving is, and why it is so difficult. He discusses standards that are available for guidance, and talks about the significant problems posed by technology obsolescence. Dr. Dollar discusses in depth the important considerations of storage media, file formats, and metadata for long term preservation and access. He cites case studies of currently operating digital archiving programs, and talks about new initiatives to watch. Finally, he talks about analog alternatives or backups as part of a long term digital archiving strategy, and wraps up the workshop with a summary of the key points.

Dr. Dollar based this workshop primarily on material contained in his book Authentic Electronic Records: Strategies for Long-Term Access, Cohasset Associates, 1999. He is in the process of updating this book, and expects a new edition to be published in May, 2006. Material is also drawn from the ISO Technical Report 18492:2005, *Long-term preservation of electronic document-based information*, a discussion draft copy of which was provided to class attendees.

Digital objects contain three attributes: the *physical*, the *logical*, and the *conceptual*. Physically, digital objects are made up of a string of binary signals recorded on a storage medium; they have no meaning by themselves. Operating software must be employed to provide a logical organization to the binary data and recognize it as a logical object based on data type. Finally, application software provides a conceptual meaning to the data objects, rendering data in human understandable form, and giving it content, context, and structure. All of these attributes must be considered when planning for digital archiving and access.

Digital archiving incorporates many activities and considerations. Electronic records must be *protected* from loss, alteration, and corruption. Their *accessibility* must be assured across organizational boundaries and across multiple technology changes and environments. Future users must be able to *use* the records in multiple ways and for many purposes, while retaining the record's meaning and authenticity. These goals must be accomplished despite ongoing changes over time in recording media, operating systems, file types and specifications, data coding systems, and metadata. 17

The process of digital archiving is made difficult because of the digital nature of the records versus traditional physical records. Physical records are easy to see, touch, understand, and manage, compared to digital records which require hardware and software to give them their logical meaning and interpretation, and provide them storage and retrievability. Electronic records require a software "interpreter" to make them understandable to humans. They are dependent on both the system operating software that makes the computer function and the application software, and they require the user to have a computer to use them. 24 Digital records can easily be rendered unusable by technological obsolescence, which is inevitable and irreversible.

The International Organization of Standards (ISO) has published several standards related to long-term preservation of electronic records and data. In particular, ISO 15489, Part 1 and 2, *Records Management,* provides the framework for an effective records management program. ISO 14721:2003, *Open archival information system – Reference model,* describes a high-level model for any electronic records repository. It sets standards for processes of data ingest, archival storage, data management, preservation planning, and access. 32 The OAIS model establishes a shared view of requirements that can lead to an interoperable network of digital archives, a key component in grid computing. 36 43 ISO 18492:2005, *Document management applications – Long-term preservation of electronic document-based information,* provides methodological guidance for the long-term preservation and retrieval of authentic electronic document-based information, when the retention period exceeds the expected life of the technology used to create and maintain the information. It sets long-term preservation goals to ensure information is readable, intelligible, identifiable, retrievable, understandable, and authentic. 46 49-50

Digital preservation requires that we deal with problems caused by technology obsolescence. For currently active electronic records, this will involve media renewal (or "refreshing") and conversion. Media renewal is the process of reformatting or copying data to new storage media to ensure its continued readability. Conversion involves the shift from one technology environment to another, such as from one version of software to a newer version, or from one software (e.g. Word Perfect) to another (Word), while maintaining the essential qualities of the electronic record. During conversion, we must maintain the processibility of the active records. 65 66

For sets of legacy records which we wish to preserve, technology obsolescence will require us to perform not only media renewal to preserve the data, but other long-term strategies as well. The Data Archaeology strategy represents the minimalist approach, in which we would keep the original data bit stream viable, and use reverse engineering in the future to devise a method to access and use that data using then-current technologies. Similar to that is the Museum Perspective, in which original hardware equipment and software versions are saved in operational condition, to be able to utilize legacy data. The Jet Propulsion Laboratory and the Washington State Digital Archives have taken this approach. For some types of evidential/informational records, Viewer Technology may provide access to images of records, without providing full functionality.

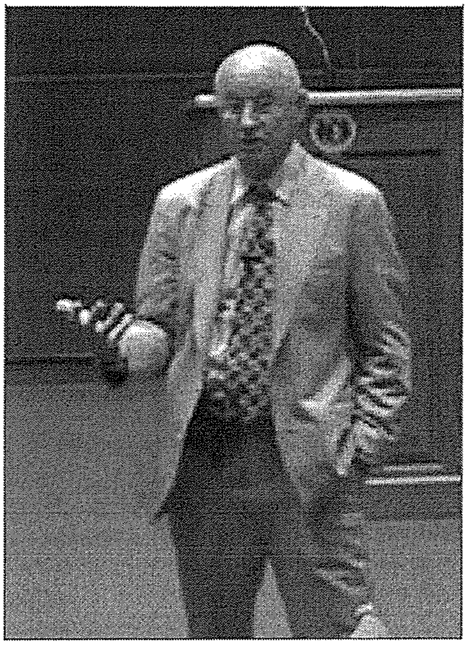
Emulation is the process of using today's computers and software to create a replica of another computer with such fidelity that it can operate in place of the other computer. Dr. Dollar discussed a number of projects designed to demonstrate the feasibility of emulation to provide access to legacy records.

Migration is an essential component of a digital preservation program. It's purpose is to ensure usable and trustworthy electronic records for as long as necessary without regard for the computer technology platform. It presumes that the bit stream remains readable through media refreshment 92 and, whenever possible, involves converting electronic records to technology neutral file formats. It should provide backward compatibility and should preserve the processibility of records. Risks associated with migration include possible alteration of the "look and feel" of records, possible loss of some data values, potential to introduce errors without good quality control, difficulty and cost of migrating complex interactive digital records, and the likelihood that the process will be never ending. Past migration efforts have shown that projects usually take longer and cost more than planned. 97

When determining the appropriate storage media for large quantities of electronic records, one must consider the speed (data transfer rate) of the selected medium, as well as its cost, capacity, and durability. 104 112 120 After discussing each storage medium in depth, Dr. Dollar concludes that magnetic media is more robust than optical, that magnetic tape holds advantages over "spinning disk" storage, and that a high data transfer rate is a vital consideration for storage and migration of huge quantities of electronic data.

- File formats tell the operating system how to interpret the 0s and 1s that comprise the electronic file. They specify the internal logical arrangement of data within digital objects, and provide special instructions such as compression algorithms. Formats also provide information understood by specific application software.

Two considerations when determining the file format for preserving electronic records include which format to use for specific information content, and whether to choose proprietary or non-proprietary formats. There are several types of electronic files, each of which have multiple formats from which to choose. Types of files include text, vector graphics, graphic images, compressed graphic images, databases, video, and audio, among others. The



concern regarding proprietary formats is that the owner of the format may restrict access to the format, or possibly go out of business and not be able to support the format in the future. Non-proprietary, open-source, widely-used formats provide a higher degree of possibility that the format will continue into the future.

An ideal file format, from a preservation point of view would have these properties:

- Device independence, without regard to the hardware/software platform
- Self-contained, containing all the resources necessary for rendering
- Self-documenting, containing its own description
- Transparency, capable of direct analysis with basic tools
- Absence of technical protection mechanisms, such as encryption, passwords, etc.
- Disclosure, with an authoritative specification publicly available
- Adoption, with widespread use being the best deterrent to obsolescence.

The recently approved PDF/A file format standard, ISO 19005, specifies how to use the Portable Document Format (PDF) 1.4 for long-term preservation of documents (/Archives). It addresses three primary issues: defining a file format that preserves the static visual appearance of electronic documents over time, providing a framework for recording metadata about electronic documents, and providing a framework for defining the logical structure and semantic properties of electronic documents. 143

In summarizing the file format discussion, Dr. Dollar recommends: 161

- Choose file formats based on recordkeeping requirements, such as integrity and processibility
- Avoid proprietary single vendor products
- Use main stream technology products
- Require transferability functionality to facilitate migration
- Consider XML, PDF, and PDF/A as good choices.

Metadata for electronic records must be captured which provides technical, business, and contextual information about the records. Technical information includes data about the creation and use of the record, the software application, and the file formats. Business information includes applicable business rules, integrity rules, and access/authorization rights. Contextual information describes "who, what, when, why", the linkage between and among records, preservation information, and offers an audit trail. Metadata is best captured at the creation or receipt of the record. While the system can provide much metadata, often users are required to key in small to large amounts of metadata.

There are presently only three operational digital archives. 170 The OCLC Digital Archives is a fee-based repository service for libraries and other institutions. Institutions can transfer electronic items to OCLC, which will preserve them and provide on-line reference services. DSpace is a digital repository system that captures, stores, indexes, preserves, and redistributes digital research materials. It is designed for academic library repositories, and requires customization to accommodate archives.

The Washington State Digital Archives is the only operating state digital archives. Planning began in 1999 and the facility opened in 2004, at an initial cost of $14.8 million. The concept is based on a well-developed feasibility study, and identifies state agency partners in terms of their level of technological sophistication and ability to transfer archival records in appropriate original formats. The project benefits from funding from a $1.00 recording fee on all filing transactions and additional support from Microsoft, a Washington corporation, and may not be easily duplicated in other states.

Besides the Washington Digital Archives, other projects under development include a demonstration project being undertaken by the Georgia State Archives with NHPRC funding, a collaboration project between the Smithsonian

Institution and Rockefeller Archives Center, and the National Archives and Records Administration's large-scale Electronic Records Archives project.

For records that must be kept long-term, consideration should be given to capturing them in microfilm or other analog format (paper) as well as digital form. The Digital Archive Writer from Kodak produces black and white microfilm from document images, and other equipment can produce color and larger-sized microfilm. A new technology, Datasurance, captures digital images in a non-proprietary 2-D barcode format, and encloses human-readable information on how to decode the barcodes, and incorporates it all on microfilm. When decoded, the barcodes recreate the original digital image. Microfilm capture should be considered for records where *reproducibility*, rather than *processibility*, will satisfy your regulatory compliance, business needs, and historical accountability.

Any organization preserving digital records must prepare a mission statement to define its purpose. It needs to define its preservation policy, describing how the mission will be carried out and specifying what activities will be done (and not done) in various circumstances. In determining the strategies to adopt, the organization looks at the convergence of available technology with its policies, as well as applicable published standards to guide them. 215 It must then identify existing best practices which it can adopt.

The threshold issues in digital preservation are to keep digital records readable, and ensure their integrity and trustworthiness over time. We cannot try to preserve everything, we must not substitute quick fixes in lieu of long-term solutions, and we should not implement technologies that are in the fringe of the marketplace.